PCT WORLD INTELLECTUAL PROPERTY (International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THI

WO 9605689A2

(51) International Patent Classification 6 : H04N

A2

(11) International Publication Number: WO 96/05689

(43) International Publication Date: 22 February 1996 (22.02.96)

(21) International Application Number: PCT/US95/09946

(22) International Filing Date: 4 August 1995 (04.08.95)

(30) Priority Data: 110573 4 August 1994 (04.08.94) IL

(71) Applicant (*for all designated States except US*): ORAD, INC. [US/US]; c/o Law Offices of Morse Geller, Suite 202, 116-16 Queens Boulevard, Forest Hills, NY 11375 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): TAMIR, Michael [IL/IL]; 13 Beit Zuri Street, 69122 Tel Aviv (IL). SHARIR, Avi [IL/IL]; 21 Ani Ma'amin Street, 67727 Ramat Hasharon (IL).

(74) Agents: GALLOWAY, Peter, D.; Ladas & Parry, 26 West 61st Street, New York, NY 10023 (US) et al.

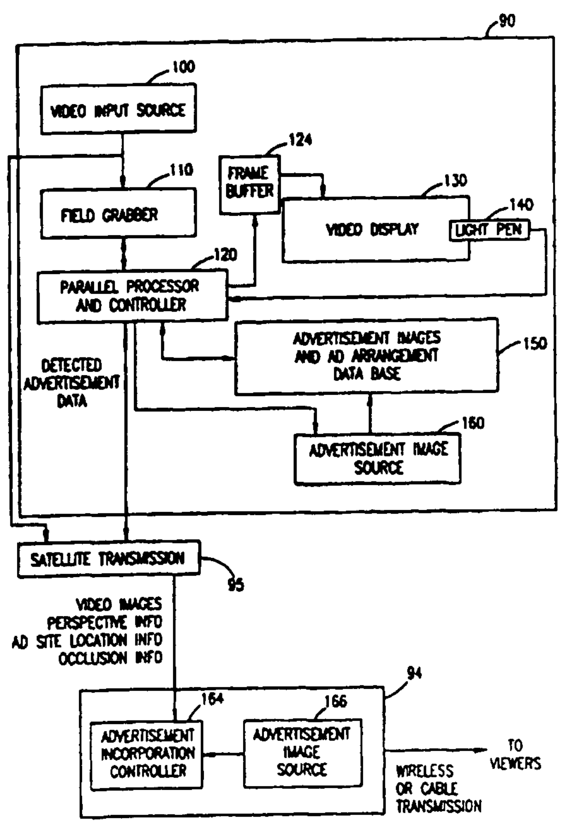
(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: APPARATUS AND METHOD FOR VIDEO BROADCASTING

(57) Abstract

This invention discloses an apparatus for replacing a portion of each of a sequence of existing images with a new image, the apparatus comprises a frame grabber operative to grab a sequence of frames respectively representing the sequence of existing images, a localizer operative to detect at least one site within each existing image at which the new image is to be incorporated, a perspective transformer operative to detect the perspective at which the site is imaged and a transmitter operative to transmit to each of a plurality of remote locations, for each frame the existing image represented in the frame, the coordinates of the site, and the perspective at which the site is imaged.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

APPARATUS AND METHOD FOR VIDEO BROADCASTING

FIELD OF THE INVENTION

The present invention relates to video broadcasting systems.

BACKGROUND OF THE INVENTION

International sports events or other spectacles generally draw the interest and attention of spectators in many countries. For example, the Olympics, Superbowl, World Cup, major basketball and soccer games and auto races fit into this category. Such events are generally broadcast live by video to a large international audience. The locale in which these events take place, such as stadiums or courts, provide advertising space all around in the form of signs, posters or other displays on fences and billboards, and in fact on any unoccupied space suitably located, including sections of the playing field.

Due to the nature of the displays, which are mostly in the form of printed matter, they are not changed too frequently and remain at least for a day, or a series or a whole season, and are directed mostly at local audiences. In cases where two teams from different countries play each other, the advertisements are occasionally arranged so that one side of the stadium contains advertisements directed to audiences in one country, while the other side has advertisements directed to the spectators in the other country.

The video cameras in these instances film the event from opposite sides of the stadium for their respective audiences. This of course is logistically

complicated and limits the angle from which the events can be seen in either of the countries represented in the game.

Another limitation to present methods of advertising is the stringent safety requirements for positioning the billboards, so as not to interfere with the game, nor disturb the view of the spectators in the stadium, nor pose a danger to the players. The displays must not be too close to the actual field of action, so as not to distract the players.

A most serious drawback of the present system for advertising at major world sports events is the fact that although the event is televised live throughout the world, the actual physical advertisements in the stadium, because of their broad international exposure, can only cater to products having a world market.

Local advertisers can only make use of such world-class televised events by locally superimposing messages on the TV screen, or by interrupting the real time of the event.

Another drawback of the existing system is that over long time periods, due to the scanning of the TV camera, the signs appear too blurred to be read by the TV viewers. On many other occasions, only part of the sign is visible to the TV viewers and the sign cannot be read.

In some applications, the requirement for computer resources is very high, on the order of magnitude of 100 BOPS (Billion Operations Per Second). To achieve this level of performance, multiple processing is employed. Many parallel-processing systems of different sizes and configurations have been developed. As the size, hardware complexity, and programming diversity of parallel systems continue to evolve, the range of alternatives for implementing a parallel task on these systems grows. There exist 2 types of models of parallelism: SIMD (Single Instruction Multi-

ple Data) machines and MIMD (Multiple Instruction Multiple Data) machines. Fig. 1A is a schematic diagram of a conventional SIMD machine.

SIMD machines contain multiple processors connected to their own memory. The PE (processing element) is a processor/memory pair. The control unit broadcasts instructions to processors. All active PEs execute the same instruction synchronously in lockstep on their own data. All machines run a single program and a single control thread (process). Various SIMD devices have been developed in academic institutions including the AMT DAP, CLIP-4, Connection Machine, Maspar MP-1 and MPP.

The MIMD machine schematically illustrated in Fig. 1B, has multiple processors connected to their own memory. The PE is a processor/memory pair. Each PE has its own instructions. PE's execute local programs on local data. All machines run multiple different programs and multiple threads of control. Examples of MIMD devices are: BBN Butterfly, Cedar, CM-5, IBM RP3, Intel Cube, Ncube, NYU Ultracomputer.

Systems which have been built to accommodate mixed MIMD/SIMD machines are the PASM in Purdure University, and Opsila.

Descriptions of parallel computer architecture in connection with image processing can be found in the following publications:

Siegal H.J. Interconnection Networks for Large-Scale Parallel Processing Theory and Case Studies, Second Edition, McGraw-Hill, New York 1990.

The following publications describe parallel solutions for image processing including MIMD and/or SIMD:

1) U.S. Patent 5,212,777 to Balmer et al describes a single MIMD chip, and does not describe in any detail a multi-chip configuration architecture.

2) Great Britain Patent No. 2250362, to Mitsubishi, describes a homogeneous system architecture based on exploiting relative simple processing elements which do not provide sufficient flexibility for graphics and video processing.

3) U.S. 4,873,626 to Gifford describes rigid selection of SIMD mode in a general case of MIMD programming mode. An I/O block for real time video processing is not described.

4) Published European Application 564847 (93104154) to IBM describes a system which is not intended for general purpose image processing. The system has only a single or dual processor array.

The following publication, the disclosure of which is incorporated herein by reference, describes Gaussian edge detection:

J.F. Canny, "A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, pp. 679-698, November, 1986.

PCT Patent Application No. WO 95/10919 titled "A method and Apparatus for Detecting, Identifying and Incorporating Advertisements in a Video" assigned to the present applicant.

SUMMARY OF THE INVENTION

The present invention relates to methods and apparatus for processing and controlling image processing using an improved parallel architecture machine for image processing and graphics. The present invention seeks to provide an improved hardware architecture which is particularly suited for parallel video and graphics processing.

Hardware parallel processors are commonly used in academic research and some commercial companies for solving image processing and vision tasks. The need for real-time in processing a video signal has prompted use of replicate hardware to allow concurrent execution. Due to the need for concurrent execution, or parallel processing, sequential processing methods may be reformulated to use parallelism.

The present invention seeks to provide an improved video broadcasting system in which an advertisement or other frame portion within a video sequence is replaced such that information identifying the coordinates of the advertisement site, the perspective at which the site is imaged, and the pixel data of occluding objects are transmitted to a remote location together with the video sequence itself. At the remote location, a replacing advertisement or other image is stored which may be specific to that location. The replacing advertisement is incorporated into the received video sequence at the appropriate site, with the appropriate scale and perspective, and properly overlaid with the occluding objects.

The present invention relates to a reconfigura-

ble parallel processing architecture machine and a new way for self-controlling the machine. The machine is a multiprocessor system which is capable of mixed-mode parallelism. It can operate in either SIMD or MIMD mode parallelism and can dynamically switch between modes. In addition, it can be partitioned into independent or communicating submachines, the architectures of which are similar to that of the original machine. Furthermore, the system uses a flexible multistage crossbar interconnection network between its processors. Furthermore, it enables any data which is to be broadcast (such as video digitized data) to be distributed to any set of processors in a controlled manner without blocking the operation of any part of the machine. Considerable variation can be accomodated as to the number of processors, the type of protocols, and the interconnection structure.

The mapping of processing routines onto the architecture shown and described herein is done using semi-automatic tools that enable simulating and running the processing routine on the invented machine. The processing routines were developed using the mixed-mode parallelism.

In the illustrated embodiment, multiple TI MVP 320C80 chips form an adaptable extendible architecture that enables the user to extend to any number of MVP's (multimedia video processors). Each MVP is a 3 billion instruction machine MIMD/SIMD VLSI chip. Each MVP is connected to its own memory blocks and connected to other MVP's by an interconnection network based on a crossbar network.

It is appreciated that the PE (processing element) forming part of the invention described herein need not be the TI MVP and that this specific implementation is given only by way of example.

Preferably, all processors are identical

in flexibility, so that a set of PE's may be selected without considering whether the task is SIMD in nature or MIMD in nature.

There is thus provided in accordance with a preferred embodiment of the present invention apparatus for replacing a portion of each of a sequence of existing images with a new image, the apparatus including a frame grabber operative to grab a sequence of frames respectively representing the sequence of existing images, a localizer operative to detect at least one site within each existing image at which the new image is to be incorporated, a perspective transformer operative to detect the perspective at which the site is imaged, and a transmitter operative to transmit to each of a plurality of remote locations, for each frame the existing image represented in the frame, the coordinates of the site, and the perspective at which the site is imaged.

There is further provided in accordance with a preferred embodiment of the present invention a method for replacing a portion of each of a sequence of existing images with a new image, the method including providing a new image, receiving from a remote transmitter, for each frame in a sequence of frames an existing image represented in the frame, coordinates of a site within the existing image at which the new image is to be incorporated, and a perspective transformation representing the perspective at which the site is imaged, applying the perspective transformation to the new image and texture mapping the transformed new image into each existing image at the site.

Additionally in accordance with a preferred embodiment of the present invention the site includes a background site and the apparatus includes an occlusion analyzer operative to identify foreground objects which at least partially occlude the background site wherein the transmitter also transmits, for each frame, an occlu-

sion map of the background site.

Further in accordance with a preferred embodiment of the present invention the method includes receiving from the remote transmitter an occlusion map of the background site and the texture mapping includes texture mapping the transformed new image into each existing image only at non-occluded locations within the site.

Still further in accordance with a preferred embodiment of the present invention the new image includes an advertisement.

Also in accordance with a preferred embodiment of the present invention each existing image includes an advertisement.

Additionally in accordance with a preferred embodiment of the present invention the apparatus includes an existing image memory operative to store an existing image and the localizer includes an image identifier operative to compare the site to the stored existing image and the transmitter is also operative to transmit, for each site, a label identifying the existing image found at the site.

Further in accordance with a preferred embodiment of the present invention the method includes receiving from the remote transmitter a label identifying the existing image found at the site, and selecting the new image according to the label.

There is further provided in accordance with a preferred embodiment of the present invention real-time video image processing apparatus operative to import and process video data from a video data source in real time, the apparatus including a first plurality of video data input/output devices, a second plurality of interconnected MIMD devices each including an array of MIMD units and at least one interconnecting bus, at least one broadcasting bus interconnecting the first plurality of video data

input/output devices and at least some of the second plurality of interconnected MIMD devices defining at least one broadcasting channel from the video data input/output devices to the MIMD devices, at least one communication bus interconnecting at least some of the second plurality of interconnected MIMD devices.

Additionally in accordance with a preferred embodiment of the present invention the at least one broadcasting bus includes a plurality of broadcasting busses.

There is also provided in accordance with a preferred embodiment of the present invention a method for controlling pipelined performance of a multi-step task by real-time video image processing apparatus, the method including providing a first plurality of video data input/output devices, a second plurality of interconnected MIMD devices each including an array of MIMD units and at least one interconnecting bus, at least one broadcasting bus interconnecting the first plurality of video data input/output devices and at least some of the second plurality of interconnected MIMD devices defining at least one broadcasting channel from the video data input/output devices to the MIMD devices, at least one communication bus interconnecting at least some of the second plurality of interconnected MIMD devices, and a library of image processing primitives, receiving a user-selected sequence including at least some of the image processing primitives in the library, constructing a pipeline to carry out the user-selected sequence which efficiently utilizes available resources, controlling the pipelined performance of the user-selected sequence.

Still further in accordance with a preferred embodiment of the present invention the method includes selecting a further pipeline depth in accordance with intermediate results. There is yet further provided in accordance with a preferred embodiment of the

present invention system control provided by a hierarchical distributed program entity provided on each PE. Functional management of the system is carried out on a system-selected number of PEs. The control software preferably includes the following functional components:

a. a real-time executive, which is a library of low level tools which combine and encapsulate the following three basic parts:

i. The mechanism of transferring data and control between PEs;

ii. Arbitration; and

iii. Support of various control structures such as semaphores and messages; and

b. a job planner, also termed herein a system task scheduler, which is operative to adapt system configuration and job distribution between PEs in order to facilitate data flow, for example by pipeline distribution of tasks between PEs.

BRIEF DESCRIPTION OF THE DRAWINGS AND APPENDICES

The present invention will be understood and appreciated from the following detailed description, taken in conjunction with the drawings and appendices in which:

Fig. 1A is a prior art simplified block diagram of a conventional SIMD device;

Fig. 1B is a prior art simplified block diagram of a conventional MIMD device;

Fig. 2 is a simplified block diagram of video broadcasting apparatus constructed and operative in accordance with a preferred embodiment of the present invention;

Fig. 3 is a simplified block diagram of the parallel processor and controller of Fig. 2;

Fig. 4 is a simplified flowchart of a preferred method of operation of the parallel processor and controller of Fig. 2, when only a single advertisement site is to be identified and only a single advertisement is to be incorporated at that site;

Fig. 5 is a simplified flowchart of a preferred method of operation of the parallel processor and controller of Fig. 2, when a plurality of advertisement sites is to be identified and a corresponding plurality of advertisements, which may or may not differ in content, is to be incorporated at those sites;

Fig. 6 is a simplified flowchart of a preferred method for performing the segmentation step of Figs. 4 and 5;

Fig. 7 is a simplified flowchart of a preferred model matching method for performing the advertisement content identification step of Figs. 4 and 5;

Fig. 8 is a simplified flowchart of a preferred method for performing the localization step of Figs. 4 and 5;

Fig. 9 is a simplified flowchart of a preferred method for performing the tracking step of Figs. 4 and 5;

Fig. 10 is a simplified flowchart of a preferred method for performing the occlusion analysis step of Figs. 4 and 5;

Fig. 11 is a simplified flowchart of a preferred method of operation for the advertisement incorporation controller of Fig. 2;

Fig. 12 is a simplified flowchart of a preferred method for detecting and tracking moving objects of central interest;

Fig. 13 is a high-level schematic block diagram of a sample real-time implementation of the parallel processor and controller of Fig. 2 including 10 boards (also termed herein "multi-MVP blocks" or MMB's) of which 9 are identical (MMB1 - MMB9) and the tenth, the input-output MMB (MMB0), also termed herein "MIOB" or "multi-input/output block", is typically different and implements the field grabber and frame buffer of Fig. 2;

Fig. 14 is a schematic block diagram of an individual one of the 9 identical MMB's (multi MVP boards) of Fig. 13, the individual MMB including 9 identical SMB's (single MVP blocks) of which one serves as a master and 8 serve as slaves with optional software based reconfiguration;

Fig. 15 is a schematic block diagram of an individual one of the SMB's of Fig. 14; and

Fig. 16 is a schematic block diagram of the input-output MMB, MMB0 of Fig. 13;

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Fig. 2 is a simplified block diagram of video broadcasting apparatus constructed and operative in accordance with a preferred embodiment of the present invention.

The apparatus of Fig. 2 includes an advertisement processing center 90 typically located in physical proximity to the video capturing equipment and a plurality of remote broadcasting stations 94 of which only one is shown. The scene being telecasted, such as a sports event, includes at least one location on which a new image, such as an advertisement, is to be superimposed. The advertisement processing center preferably generates and transmits to each remote broadcasting station 94 all information needed to superimpose a new image onto the advertisement, such as the location of the site at which the advertisement is to appear, its scale, the perspective at which the site is imaged and the locations of foreground objects occluding the site. At each remote station 94, the new advertisements to be incorporated are stored and these advertisement are easily incorporated into the arriving video frames, using the information generated by the advertisement processing center 90.

A particular advantage of the apparatus of Fig. 2 is that the function of analysis of the existing video frames is separated from the function of incorporating the new advertisement into the frame. The information employed to superimpose new images onto the advertisement is generated only once and is then employed by all of the broadcasting stations to facilitate incorporation of the new advertisement into the frame.

The advertisement processing center 90 includes a video input source 100, such as a video camera, video

cassette, broadcast, video disk, or cable transmission, which is connected, via a suitable connector, with a field grabber 110, preferably, or alternatively with a frame grabber. Henceforth, use of the term "field grabber" is intended to include frame grabbers.

The field grabber 110 provides grabbed and digitized fields to a parallel processor and controller 120, (described in more detail below with reference to Fig. 3) which automatically detects, identifies, and localizes a given advertisement in the field of view, and is preferably associated, via a frame buffer 124, with a video display 130 which provides an interactive indication to a user of advertisement site detection operations of the system. Preferably a light pen 140 is associated with the video display 130 or alternatively the video display 130 comprises a touch screen.

According to an alternative embodiment of the present invention, the system receives an indication from a user of the presence in the field of view of one or more advertisements to be replaced and of the location/s thereof. The user input may, for example, be provided by means of a light pen 140. The indication provided by the user may comprise a single indication of an interior location of the advertisement, such as the approximate center of the advertisement or may comprise two or four indications of two opposite vertices or all four vertices, respectively, of an advertisement to be replaced.

Optionally, the user also provides an indication of the contents of the advertisement. For example, a menu of captions identifying advertisements to be replaced, may be provided on the video display 130 adjacent or overlaying a display of the playing field and the user can employ the light pen to identify the appropriate caption.

An advertisement images and advertisement arrangement database 150 is preferably provided which may

be stored in any suitable type of memory such as computer memory or secondary memory, such as a hard disk. The advertisement image and arrangement database 150 stores an indication of the arrangement of a plurality of advertisements to be replaced, if the arrangement is known ahead of time. Typically, the indication of the arrangement does not include an indication of the location of each advertisement relative to the playing field, but instead includes an indication of the order in which the advertisements to be replaced will be arranged in the field. For example, a sequence of 20 side-by-side advertisements may be arranged around three sides of a playing field. The database 150 may then include an indication of the sequence in which the advertisements are arranged.

The database 150 also preferably stores images of the advertisements to be replaced so that these can be recognized.

Advertisement images in the database 150 may be provided by field grabber 110 or from any suitable advertisement image source 160, such as but not limited to an image generating unit such as an image processing workstation, a scanner or other color reading device, any type of storage device, such as a hard disk, a CD ROM driver, or a communication link to any of the above.

The output of parallel processor and controller 120, also termed herein "the auxiliary output", comprises, for each frame, an indication of the locations at which advertisements are to be incorporated, an indication of the perspective at which each advertisement is to be incorporated and an indication of the locations of foreground objects, if any, occluding each advertisement. The auxiliary output may alternatively comprise the scale, location and perspective transformation parameters of other identified landmarks in the video image.

The video and auxiliary output of the system may be provided via a suitable connector to suitable

equipment for providing satellite or other transmission 95 to a remote broadcasting station 94.

Any suitable video broadcasting technology may be employed, such as via satellites, via RF transmission or via cable optics systems.

The volume of the auxiliary information regarding the video original, is small relative to the video original itself. The auxiliary information may be sent on the audio channel or alternatively may be sent on a separate channel but synchronized to the channel over which the video original is being sent by means of a common time code. Alternatively, the bandwidth of the link may be slightly increased so as to enable the auxiliary information and the video original to be sent on the same video channel.

The broadcasting station 94 includes conventional broadcasting equipment (not shown) for wireless or cable transmission to viewers and, additionally, an advertisement incorporation controller 164 accessing an advertisement image source 166. The image source 166 may comprise a hard disk storing a plurality of advertisement images, typically still images, to be incorporated into the image of the playing field, either replacing an existing advertisement or in a location not presently occupied by an advertisement.

The advertisement incorporation controller 164 is operative to receive the video frames and the auxiliary information from the advertisement processing center 90 and to incorporate the advertisement image or images of image source 166 into the appropriate locations in each video frame, at the appropriate perspectives and with the appropriate portions thereof removed such that occlusion of the original location by a foreground object is maintained for the incorporated advertisement image. Optionally, a predetermined advertisement incorporation schedule is followed such that different advertisements

are incorporated at the same location, each for a predetermined length of time within the game or upon a predetermined occurrence within the game, such as a goal.

A preferred method of operation for controller 164 is described below with reference to Fig. 11.

Fig. 3 is a simplified block diagram of the parallel processor and controller 120 of Fig. 2. The parallel processor/controller 120 preferably includes an advertisement site detection/content identification unit 170, a plurality of parallel tracking modules 180, an occlusion analysis unit 190, and a controller 210.

The advertisement site detection/content identification unit 170 of Fig. 3 may be implemented based on a suitable plurality of suitable image processing boards. A specially designed coprocessor is preferably added to these boards to perform the segmentation task. The image processing boards are programmed based on the advertisement site detection and content identification methods of Figs. 6 and 7. Each of parallel tracking modules 180 may be implemented based on one or more image processing boards. The image processing boards are programmed for parallel operation based on the tracking method of Fig. 9. The occlusion analysis unit 190 may also be based on one or more multi-DSP (Digital Signal Processing) boards, programmed based on the occlusion analysis of Fig. 10. Controller 210 may, for example, comprise a Silicon Graphics Indy Workstation, programmed based on the control method of Figs. 4 - 5.

Fig. 4 is a simplified flowchart of a preferred method of operation of the parallel processor and controller 120 of Fig. 2, when only a single advertisement site is to be identified.

Fig. 5 is a simplified flowchart of a preferred method of operation of the parallel processor and controller 120 of Fig. 2, when a plurality of advertisement sites is to be identified.

The method of Fig. 5 typically includes the following steps, which are similar to the steps of Fig. 4 which are therefore not described separately for brevity:

STEP 290: A digitized video field is received from the field grabber 110 of Fig. 2.

STEP 300: A decision is made as to whether or not at least one advertisement in the current field was also present in the previous field (and televised by the same camera). If so, the current field is termed a "consecutive" field and the segmentation, content identification and localization steps 320, 330 and 340 preferably are replaced only by a tracking step 310. If not, the current field is termed a "new" field.

If the field is a "consecutive" field, the plurality of advertisements is tracked (step 310), based on at least one advertisement which was present in a previous field, since the present field is a "consecutive" field.

If the field is a "new" field, the advertisement site at which an advertisement is to be incorporated is identified in steps 320, 330 and 340. A loop is performed for each advertisement from among the plurality of advertisements to be processed. Preferably, the segmentation and content identification steps 320 and 330 are performed only for the first advertisement processed.

In step 320, a pair of generally parallel lines is typically detected and the image of the field is segmented. Specifically, the portion of the field located within the two detected parallel lines, which typically correspond to the top and bottom boundaries of a sequence of advertisements, is segmented from the remaining portion of the field.

Typically, the segmentation step 320 is operative to segment advertisements regardless of: the zoom state of the imaging camera lens, the location of the advertisement in the field of view (video field), the

angular orientation of the imaging camera relative to the ground and the location of the TV camera.

The segmentation step 320 is typically operative to identify an empty or occupied advertisement site, such as but not limited to any of the following, separately or in any combination:

a. Geometrical attributes of the advertisement's boundary such as substantially parallel top and bottom boundaries or such as four vertices arranged in a substantially rectangular configuration;

b. A color or a combination of colors or a color pattern, which is known in advance to be present in the advertisement image.

c. The spatial frequencies band of the advertisement image, which is typically known in advance. Typically, the known spatial frequencies band is normalized by the height of the advertisement which may, for example, be derived by computing the distance between a pair of detected horizontal lines which are known to be the top and bottom boundaries of the advertisement sequence.

In step 330, the content of the portion between the two substantially parallel lines is matched to a stored representation of an advertisement to be replaced.

Steps 320 and 330 allow advertisement sites to be identified and the content thereof to be matched to a stored model thereof, even if cuts (transitions, typically abrupt, between the outputs of a plurality of cameras which are simultaneously imaging the sports event) occur during the sports event. Typically, at each cut, steps 320 and 330 are performed so as to identify the advertisement within the first few fields of the cut. Until the next cut occurs, the identified advertisement is typically tracked (step 310).

In step 340, the advertisement is localized at subpixel accuracy.

Finally, for each advertisement, occlusion

analysis is performed (step 350).

According to an alternative embodiment of the present invention, the segmentation and advertisement content identification steps 320 and 330 respectively may be omitted if physical landmarks identifying the locations of advertisements to be replaced whose contents is known in advance, are positioned and captured ahead of time in the playing field.

Fig. 6 is a simplified flowchart of a preferred method for performing the segmentation step 320 of Figs. 4 and 5.

The method of Fig. 6 preferably includes the following steps:

STEP 380: A new field is received and the resolution thereof is preferably reduced since the forgoing steps may be performed adequately at a lower resolution. for example, a low pass filter may be employed to reduce a 750 x 500 pixel field to 128 x 128 pixels.

STEP 390: Optionally, the low resolution image is smoothed, e.g. by median filtering or low pass filtering, so as to remove information irrelevant to the task of searching for long or substantially horizontal lines.

STEP 400: Edges and lines are detected, using any suitable edge detection method such as the Canny method, described by J.F. Canny in "A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, pp. 679-698, November, 1986.

STEP 404: The edges detected in step 400 are thinned and components thereof are connected using conventional techniques of connectivity analysis. The edges are thresholded so as to discard edges having too small a gradient.

STEP 408: The edges detected in steps 400 and 404 are compared pairwise so as to find strips, i.e. pairs of parallel or almost parallel lines which are

relatively long. If there are no such pairs, the method terminates.

STEP 412: Find the spatial frequency spectrum within each strip and reject strips whose spatial frequency contents are incompatible with the spatial frequency band expected for advertisements. Typically, the rejection criterion is such that more than one strip, such as 3 or 4 strips, remain.

STEP 416: Rank the remaining strips and select the highest ranking strip. The rank assigned to a strip depends on the probability that the strip includes advertisements. For example, the strip in the lowest location in the upper half of the field is given higher rank than strips above it, because the strips above it are more likely to be images of portions of the stadium. The lowest located strip is more likely to be the advertisements which are typically positioned below the stadium.

Strips adjacent the bottom of the field are given low rank because the advertisements would only be imaged toward the bottom of the video field if the playing field is not being shown at all, which is unlikely.

Fig. 7 is a simplified flowchart of a preferred model matching method for performing the advertisement content identification step 330 of Figs. 4 and 5. Alternatively, advertisement content identification may be provided by a user, as described above with reference to Fig. 2.

The method of Fig. 7 is preferably performed in low resolution, as described above with reference to step 380 of Fig. 6. The method of Fig. 7 preferably includes the following steps:

STEP 420: The forgoing steps 424, 430, 436, 440, 444 and 452 are performed for each almost parallel strip identified in segmentation step 320 of Figs 4 and 5.

STEP 424: The distance and angle between the two lines of each strip is computed and the scale and approximate perspective at which the strip was imaged is determined therefrom.

STEP 430: During set-up, each advertisement model is divided into a plurality of windows. Steps 436, 440 and 444 are performed for each window of each advertisement model. For example, if there are 5 models each partitioned into 6 windows, this step is performed 30 times.

STEP 436: A one-dimensional similarity search is carried out for the suitably scaled current model window k, along the current almost parallel strip. Typically, a cross-correlation function may be computed for each pixel along the current strip.

STEP 440: The cross-correlation function values obtained in step 436 are thresholded. For example, values exceeding 0.6 may be assigned the value 1 (correlation) whereas values under 0.6 may be assigned the value 0 (no correlation). The 1's are weighted, depending on the "significance" of their corresponding windows. The "significance" of each window is preferably determined during set-up such that windows containing more information are more "significant" than windows containing little information.

STEP 444: At this stage, weighted thresholded cross-correlation function values have been computed which represent the results of matching the contents of each position along the strip (e.g. of each of a plurality of windows along the strip which are spaced at a distance of a single pixel) to each window of each model advertisement known to occur within the strip.

The weighted thresholded cross-correlation function values are accumulated per all windows composing a model sign or a model strip.

STEP 452: A decision is made as to the approxi-

mate location of the sequence of advertising models, within the strip. It is appreciated that, once the location of one advertisement model has been determined, the locations of the other advertisement models in the same sequence are also determined, knowing the scale and approximate perspective of the imaged strip.

Fig. 8 is a simplified flowchart of a preferred method for performing the precise localization step 340 of Figs. 4 and 5. In Fig. 8, the advertisement model which was approximately localized by the method of Fig. 7, is localized with subpixel accuracy. Accurate localization is typically performed only for new fields. For "consecutive" fields, the advertisement's location is preferably measured by video tracking.

The method of Fig. 8 preferably includes the following steps:

STEP 460: From Fig. 7, the following information is available per advertisement detected: one location within the advertisement, such as one vertex thereof, the advertisement scale height in the image and its approximate perspective. This information is employed to compute the four vertices of each detected advertisement (sign).

STEP 464: A perspective transformation is computed which describes how to "transform" the typically rectangular model into the detected advertisement area which is typically non-rectangular due to its pose relative to the imaging camera.

STEP 468: The contents of each of a plurality of model tracking windows to which the model is divided during set up, is mapped into the video field, using the perspective transformation computed in step 464.

STEP 470: Steps 472 and 476 are performed for each of the model tracking windows.

STEP 472: The current model tracking window is

translated through a search area defined in the video field. For each position of the model tracking window within the search area, a similarity error function (like cross-correlation or absolute sum of differences) is computed. Typically, the model tracking window has 8 x 8 or 16 x 16 different positions within the search area.

STEP 476: The minimum similarity error function for the current model tracking window is found. Preferably, the minimum is found at subpixel accuracy, e.g. by fitting a two-dimensional parabola to the similarity error function generated in step 472 and computing the minimum of the parabola. This minimum corresponds to the best position, at "subpixel accuracy", for the current model tracking window within the video field.

If (STEP 480) the similarity error function minima are high for all tracking windows, i.e. none of the tracking windows can be well matched to the video field, then (STEP 482) processing of the current frame is terminated and the method of Fig. 4, from step 320 onward, is performed on the following frame.

STEP 484: Tracking windows which have a high similarity error function minimum are rejected. Typically, approximately 30 tracking windows remain.

STEP 488 is a stopping criterion determining whether or not to perform another iteration of localization by matching tracking windows. Typically, if the tracking windows' centers are found to converge, relative to the centers identified in the last iteration, the process is terminated. Otherwise, the method returns to step 464.

STEP 490: Once the tracking window locations have converged, the perspective transformation between the image's advertisement and its model is recomputed.

Fig. 9 is a simplified flowchart of a preferred method for performing the tracking step 310 of Figs. 4 and 5. The method of Fig. 9 preferably includes

the following steps:

STEP 492: A perspective transformation is performed on the model tracking windows and the contents thereof are mapped into the video field. This step employs the system's knowledge of the location of the advertisement in the previous field and, preferably, predicted scanning speed of the camera imaging the sports event.

STEP 496: Steps 498 and 500, which may be similar to steps 472 and 476, respectively, of Fig. 8, are performed for each model tracking window.

STEPS 508 and 512 may be similar to steps 488 and 490 of Fig. 8.

STEP 510: If the window center locations do not yet converge, step 492 is redone, however, this time, the texture mapping is based upon the perspective transformation of the previous iteration.

STEP 520: The coefficients of the perspective transformation are preferably temporally smoothed, since, due to the smoothness of the camera's scanning action, it can be assumed that discontinuities are noise.

Fig. 10 is a simplified flowchart of a preferred method for performing the occlusion analysis step 350 of Figs. 4 and 5. The method of Fig. 10 preferably includes the following steps:

STEP 530: The advertisement image in the video field is subtracted from its perspective transformed model, as computed in step 512 of Fig. 9 or, for a new field, in step 490 of Fig. 8.

STEP 534: Preferably, the identity of the advertisement image and the stored advertisement is verified by inspecting the difference values computed in step 530. If the advertisement image and the stored advertisement are not identical, the current field is not processed any further. Instead, the next field is processed, starting from step 320 of Fig. 5.

STEP 538: The internal edge effects are filtered out of the difference image computed in step 530 since internal edges are assumed to be artifacts.

STEP 542: Large non-black areas in the difference image are defined to be areas of occlusion.

STEP 546: The occlusion map is preferably temporally smoothed since the process of occlusion may be assumed to be continuous.

Fig. 11 is a simplified flowchart of a preferred method for performing the advertisement incorporation step 164 of Fig. 2. The method of Fig. 11 preferably includes the following steps:

STEP 560: The resolution of the replacing advertisement model, i.e. the advertisement in memory, is adjusted to correspond to the resolution in which the advertisement to be replaced was imaged. Typically, a single advertisement model is stored in several different resolutions.

STEP 570: The replacing advertisement is transformed and texture mapped into the video field pose, using tri-linear interpolation methods. This step typically is based on the results of step 512 of Fig. 9 or, for a new field, on the results of step 490 of Fig. 8 relayed to the incorporation unit 94 via the transmission link.

STEP 580: Aliasing effects are eliminated.

STEP 584: The replacing pixels are keyed in according to an occlusion map. The values of the replacing pixels may either completely replace the existing values, or may be combined with the existing values, as by a weighted average. For example, the second alternative may be used for edge pixels whereas the first alternative may be used for middle pixels.

The applicability of the apparatus and methods described above are not limited to the detection, tracking and replacement or enhancement of advertisements. The

disclosed apparatus and methods may, for example, be used to detect and track moving objects of central interest, as shown in Fig. 12, such as focal athletes and such as balls, rackets, clubs and other sports equipment. The images of these moving objects may then be modified by adding a "trail" including an advertisement such as the logo of a manufacturer.

Fig. 13 is a high-level schematic block diagram of a sample real-time implementation of the parallel processor and controller of Fig. 2 including 10 boards (MMB's) of which 9 are identical (MMB1 - MMB9) 600 and the tenth, the input-output MMB (MMB0) 601, is typically different and implements the field grabber and frame buffer of Fig. 2.

Fig. 14 is a schematic block diagram of an individual one of the 9 identical MMB's (multi MVP boards) of Fig. 13, the individual MMB including 9 identical SMB's (single MVP blocks) of which one serves as a master and 8 serve as dynamically reconfigurable slaves.

Fig. 15 is a schematic block diagram of an individual one of the SMB's of Fig. 14.

Fig. 16 is a schematic block diagram of the input-output MMB, MMB0 of Fig. 13. The machine vision methods provided by the present invention employs a tremendous amount of and variety of parallelism. There are three processing levels in vision: low (sensory), intermediate (symbolic), and high (knowledge-based).

In addition to the vision task, the present invention provides for synthesis of a newly created image. A combination of image processing and computer graphics operations are employed. A particular feature of the apparatus shown and described herein is the capability to perform real time video image processing (vision) operations as well as graphics operations in which the image is modified as by adding an artificially created

image thereto. For example, a complex texture mapping operation is performed in which a newly computed true perspective texture mapped sign replaces an existing sign in the image.

In vision, a typical sensor comprises a camera with resolution of, for example, 720 by 560 pixels each represented by 3 color components: Red, Green and Blue. A plurality of blocks of the image are typically processed in parallel. The images from the video source tend to stream steadily into the machine which preferably requires pipeline parallelism. The present invention allows multiple sensor data to be exploited, thus providing yet another potential source of parallelism. The system shown and described herein extracts many features from a given image or set of images, such as lines, regions, texture patches, and motion parameters. These processes are preferably carried out in parallel.

The system shown and described herein is operative to add or replace a recognized object within an image in a video sequence, to/with newly created 3D objects projected to the image plane with the right perspective, lighting and blur effect.

The system shown and described herein includes three main blocks of hardware:

a. A processing block which includes a plurality of PE's (processing elements) each preferably comprising an MVP TI chip with memory and a data exchanging module operative to exchange data with the other PE's.

b. An input/output module which inputs and outputs a plurality of video signals to and from the processing block.

c. Interconnection blocks that are embedded in each PE block and each module card, and enable the exchange of data by means of crossbar switches and enabling hardware to broadcast any type of data from one PE to any set of PEs.

The three levels of vision (low, intermediate, and high) and graphics are handled on the same PE which may be implemented by the TI MVP chip. Each TI MVP chip is VLSI, structured with several individual processors all having communication links to several memories. A crossbar switch is used to establish the processor memory links. Each processor is operative to execute the same instruction at the same time (SIMD mode) or different instructions at the same time (MIMD mode).

The present invention provides an effective way of using multiple TI chips based on hierarchical architecture. In particular, each board contains one master MVP, and each set of boards includes a System Manager MVP chip, thus providing the possibility of reconfiguration of hardware and software, depending on a particular application.

Fig. 13 is an overview of a preferred architecture. The system is contained in a box or a set of boxes each containing a number of identical boards, also termed herein "MMB's" (Multiple MVP block). The system has a communication channel between nine PEs, one of which is used as a master for the other eight. Any of the PE's can be used for any type of processing. Each of the PE's is configured on an SMB (Single MVP block) which contains a full environment for 1 MVP together with memory and communication means for the crossbar block. All the SMBs on each board are interconnected by a crossbar, which provides at least 4 concurrently different channels between the PEs as well as the ability to broadcast any set of data blocks between one PE (TI-MVP) to a set of other PEs.

Each communication channel is connected on-board to shared memory. The system is preferably operative to transfer a raw video signal from the video I/O block to all the PEs without interfering with operation

of each of the PE's.

Fig. 14 is a simplified block diagram of an individual MMB. As shown, each MMB includes:

1) A video bus control block 608 which receives video data from the video I/O block in real time and broadcasts it to any of the slaves on the board. The video bus control block 608 is controlled by the SMB #9 (606).

2) A plurality of SMBs (PEs) 604 and 606, such as 9 SMB's in the illustrated embodiment. Each SMB contains a PE and has two buses 622 and 624 (Fig. 15) of which one is a slave bus 624 and the other is a master bus 622. Each of the buses is connected to a different memory block so that data can move through the slave bus and through the master bus concurrently.

3) CBT - (Crossbar transceivers) also termed herein Crossbar switches 610 which, via the master SMB (PE) 606, provide flexible channel selection. The CBT 610 includes the master bus 622 (Fig. 15). The channel is then arbitrated and the CBT 610 reconfigured such that only a set of SMBs 604 is connected. Thereby, communication between any of the SMBs 604 may be either point to point or as a broadcast operation. One CBT is connected to each SMB which allows the master bus to be connected to one of the four communication channels 611, 612, 613 and 614. When one SMB wishes to connect to a set of other SMBs, it arbitrates for a free channel. When the SMB receives the free channel, it acknowledges this to the master SMB 606 and then sends the required data packets from itself to all the other SMBs. The CBTs preferably share memory, one of the shared block 615 with shared memory block 600.

The SMB's can also communicate between themselves via the master processor which uses the master bus to provide communication between the SMBs. After a connection is formed, one PE may be selected by the

master to broadcast a block of data to the members in the channel. Alternatively, any of the SMBs may send a packet serially over the network.

Fig. 15 is a block diagram of an individual SMB. Each of the PE - MVP chips 620 is flexible enough to accommodate substantially any type of image processing and graphics processing. The control method shown and described herein includes a real time executive which facilitates flexibility such that a single process may be run on several MVPs on the same MMB and/or on more than one MMBs.

For each SMB, there is one MVP that includes four 32 bits PE integer processors and one 32 bits MP RISC (Reduced Instruction Set Computer) processor, including a floating processor. Each MVP enables operation of up to 60 concurrent RISC operations per cycle (3 billion operations per second) and 100 MFLOPS RISC. One box containing 200 MVPs may perform 600 billion operations per second and 20 billion floating operations per second.

Fig. 16 is a block diagram of the I/O 601 of Fig. 13 staging module which permits one or more sensors to input images into a video buffer 640. The buffer 640 can hold several frames for each sensor in a pipeline and broadcast any set of them to any set of MMBs. The input video connects to a block 632 which transfers the video to the global bus through a global connector 633.

The output video from the system is transferred through an output block 634 which translates digital video in the system to standard CCIR 601 D1 signals. A LAN (Local Area Network) connect block 636 enables control of the system from any host external computer, such as computer 642 (Fig. 13) through a LAN (Local Area Network) connection 638. The system is reconfigurable in its pipeline depth. In other words, each MMB

according to its task in the process stores the required number of images. In the embodiment shown and described herein, a pipeline depth of 6 is required. The block can sustain full resolution broadcast video signals at 30 frames per second on NTSC (US Video Standard) or 25 frames with PAL (European Video Standard) signalling. The operation may be performed either at field level or at frame level.

System control is provided by a hierarchical distributed program entity provided on each PE. Functional management of the system is carried out on a system-selected number of PEs. The control software preferably includes the following functional components:

a. a real-time executive, which is a library of low level tools which combine and encapsulate the following three basic parts:

i. The mechanism of transferring data and control between PEs;

ii. Arbitration; and

iii. Support of various control structures such as semaphores and messages; and

b. a job planner, also termed herein a system task scheduler, which is operative to adapt system configuration and job distribution between PEs in order to facilitate data flow, for example by pipeline distribution of tasks between PEs.

In advertisement replacement applications, the job planner selects, according to the location of and other attributes of objects in the current image, a sequence of macro operations. Each macro operation itself comprises a sequence of the low-level tools in the library. The job planner also dynamically assigns a specific PE to execute certain macro operations. For example, the number of players occluding a sign or a portion of a sign and the total occluded area may each affect the distribution of tasks between PE's. In contrast, a hard-

ware system is less adaptive and therefore less effective.

It is appreciated that the architecture shown and described herein may be extended to an almost unlimited extent.

It is appreciated that the software components of the present invention may, if desired, be implemented in ROM (read-only memory) form or be loaded into RAM (random access memory). The software components may, generally, be implemented in hardware, if desired, using conventional techniques.

It is appreciated that various features of the invention which are, for clarity, described in the contexts of separate embodiments may also be provided in combination in a single embodiment. Conversely, various features of the invention which are, for brevity, described in the context of a single embodiment may also be provided separately or in any suitable subcombination.

It will be appreciated by those skilled in the art that the invention is not limited to what has been shown and described hereinabove. Rather, the scope of the invention is defined solely by the claims which follow:

CLAIMS

1. Apparatus for replacing a portion of each of a sequence of existing images with a new image, the apparatus comprising:

a frame grabber operative to grab a sequence of frames respectively representing the sequence of existing images;

a localizer operative to detect at least one site within each existing image at which the new image is to be incorporated;

a perspective transformer operative to detect the perspective at which the site is imaged; and

a transmitter operative to transmit to each of a plurality of remote locations, for each frame:

the existing image represented in the frame,

the coordinates of the site, and

the perspective at which the site is imaged.

2. A method for replacing a portion of each of a sequence of existing images with a new image, the method comprising:

providing a new image;

receiving from a remote transmitter, for each frame in a sequence of frames:

an existing image represented in the frame,

coordinates of a site within the existing image at which the new image is to be incorporated; and

a perspective transformation representing the perspective at which the site is imaged;

applying the perspective transformation to the

new image and texture mapping the transformed new image into each existing image at the site.

3. Apparatus according to claim 1 wherein the site comprises a background site and the apparatus also comprises an occlusion analyzer operative to identify foreground objects which at least partially occlude the background site wherein the transmitter also transmits, for each frame, an occlusion map of the background site.

4. A method according to claim 2 and also comprising receiving from the remote transmitter an occlusion map of the background site and wherein said texture mapping comprises texture mapping the transformed new image into each existing image only at non-occluded locations within said site.

5. Apparatus according to claim 1 or claim 3 wherein the new image comprises an advertisement.

6. Apparatus according to claim 1 or claim 3 wherein each existing image also comprises an advertisement.

7. Apparatus according to any of claims 1, 3, 5 or 6 and also comprising an existing image memory operative to store an existing image and wherein the localizer comprises an image identifier operative to compare the site to the stored existing image and wherein the transmitter is also operative to transmit, for each site, a label identifying the existing image found at the site.

8. A method according to claim 2 or claim 4 and also comprising:

receiving from the remote transmitter a label identifying the existing image found at the site; and

selecting the new image according to said label.

9. Apparatus as claimed in claim 1 in which the localizer comprises:

multiple video data input/output devices and multiple interconnected MIMD devices each including an array of SMB units and at least one interconnecting bus; and

at least one broadcasting bus for interconnecting the video data input/output devices with selected ones of the SMB devices.

10. Apparatus according to claim 9 comprising a memory containing a library of image processing primitives, means for receiving a selected sequence including at least some of the image processing primitives in the library, a pipeline to carry out the selected sequence which efficiently utilizes available resources, and means for controlling pipelined performance of the selected sequence.

11. Real-time video image processing apparatus operative to import and process video data from a video data source in real time, the apparatus including:

a first plurality of video data input/output devices;

a second plurality of interconnected MIMD devices each including an array of MIMD units and at least one interconnecting bus;

at least one broadcasting bus interconnecting the first plurality of video data input/output devices and at least some of the second plurality of interconnected MIMD devices defining at least one broadcasting channel from the video data input/output devices to the MIMD devices; and

at least one communication bus interconnecting at least some of the second plurality of interconnected MIMD devices.

12. Apparatus according to claim 11 wherein the at least one broadcasting bus includes a plurality of broadcasting busses.

13. A method for controlling pipelined performance of a multi-step task by real-time video image processing apparatus, the method including:

providing a first plurality of video data input/output devices, a second plurality of interconnected MIMD devices each including an array of MIMD units and at least one interconnecting bus, at least one broadcasting bus interconnecting the first plurality of video data input/output devices and at least some of the second plurality of interconnected MIMD devices defining at least one broadcasting channel from the video data input/output devices to the MIMD devices, at least one communication bus interconnecting at least some of the second plurality of interconnected MIMD devices, and a library of image processing primitives, and

receiving a user-selected sequence including at least some of the image processing primitives in the library, constructing a pipeline to carry out the user-selected sequence which efficiently utilizes available resources, and controlling the pipelined performance of the user-selected sequence.

14. A method according to claim 13 and also including selecting a further pipeline depth in accordance with intermediate results.

15. Apparatus according to claim 12 wherein functional management of the system is carried out on a

system-selected number of PE's using control software which preferably includes:

a real-time executive comprising a library of low level tools which combine and encapsulate basic parts; and

a job planner which is operative to adapt system configuration and job distribution between PEs in order to facilitate data flow.

16. Apparatus according to claim 15 wherein the basic parts comprise:

a mechanism of transferring data and control between PEs;

arbitration apparatus; and

control structure support means.

17. Apparatus according to claim 15 or claim 16 wherein data flow is facilitated by pipeline distribution of tasks between PEs.

FIG. 1A
PRIOR ART

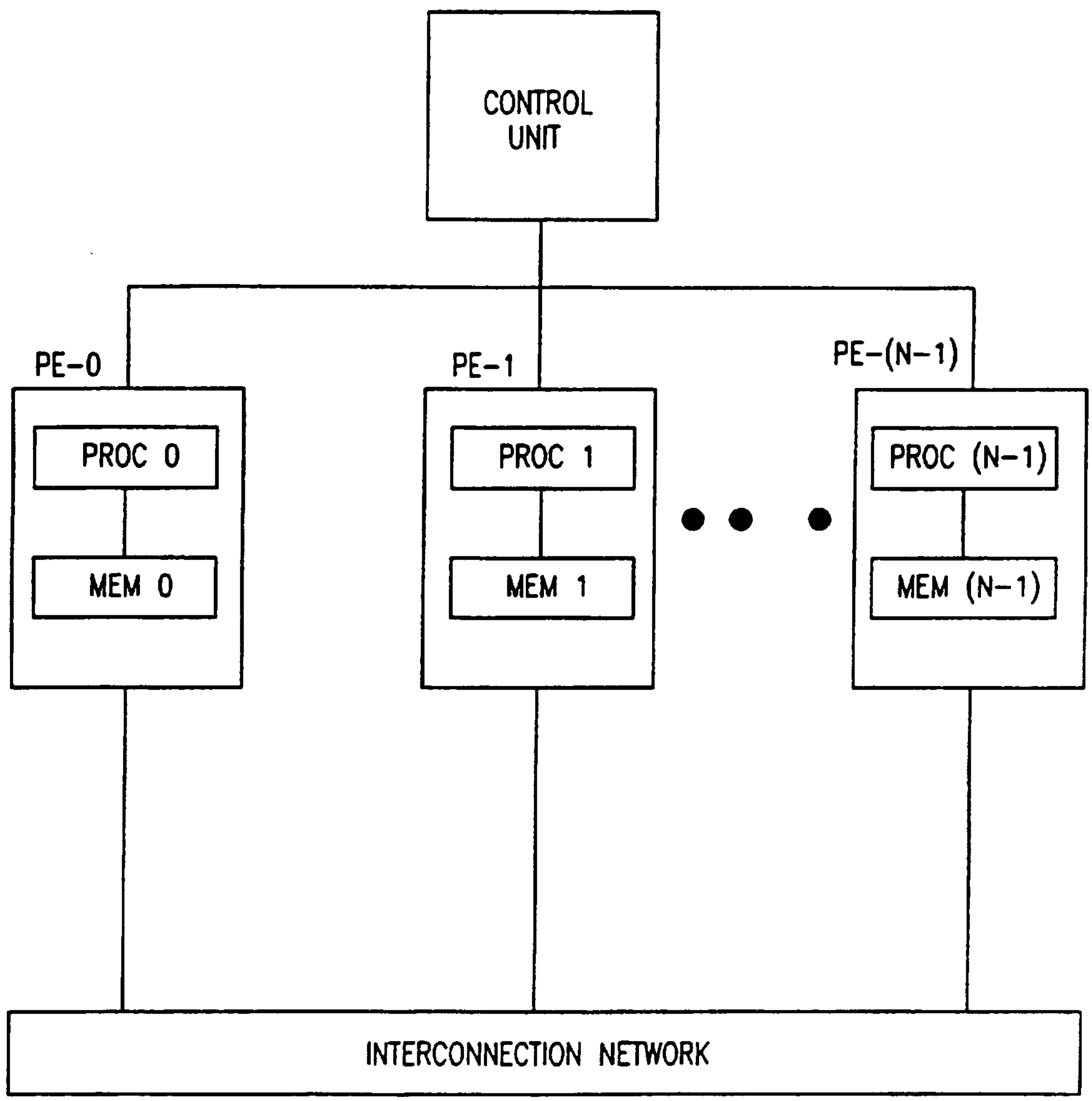
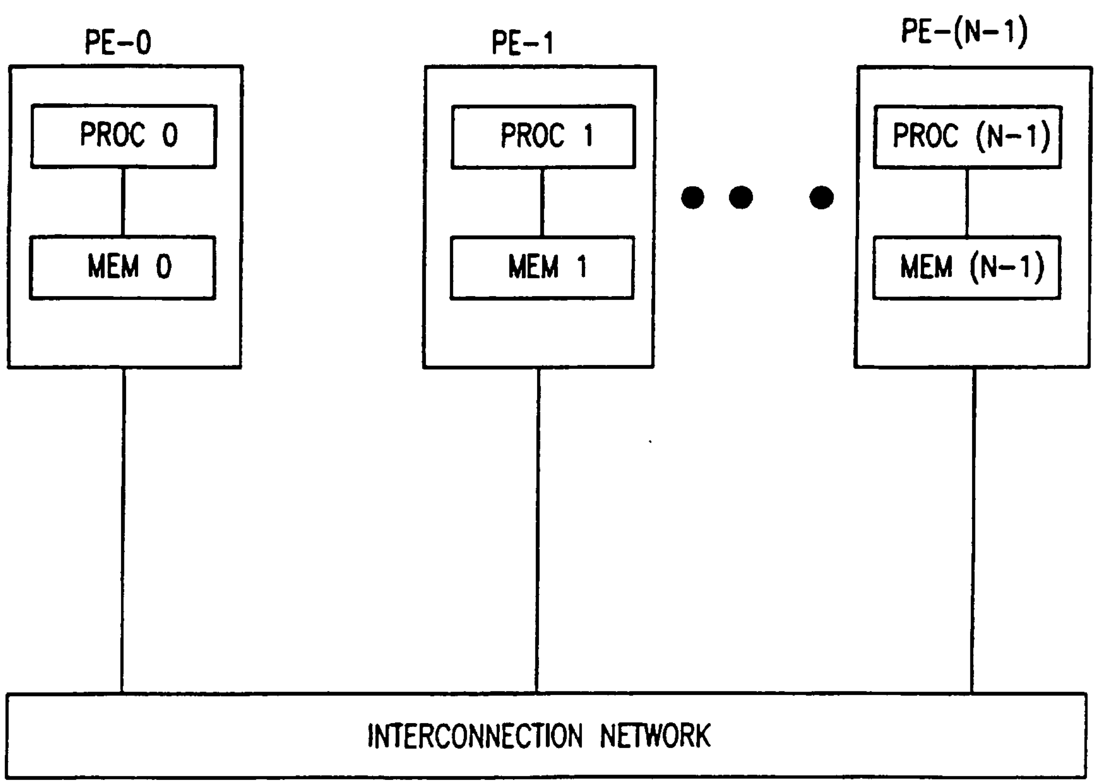


FIG. 1B
PRIOR ART



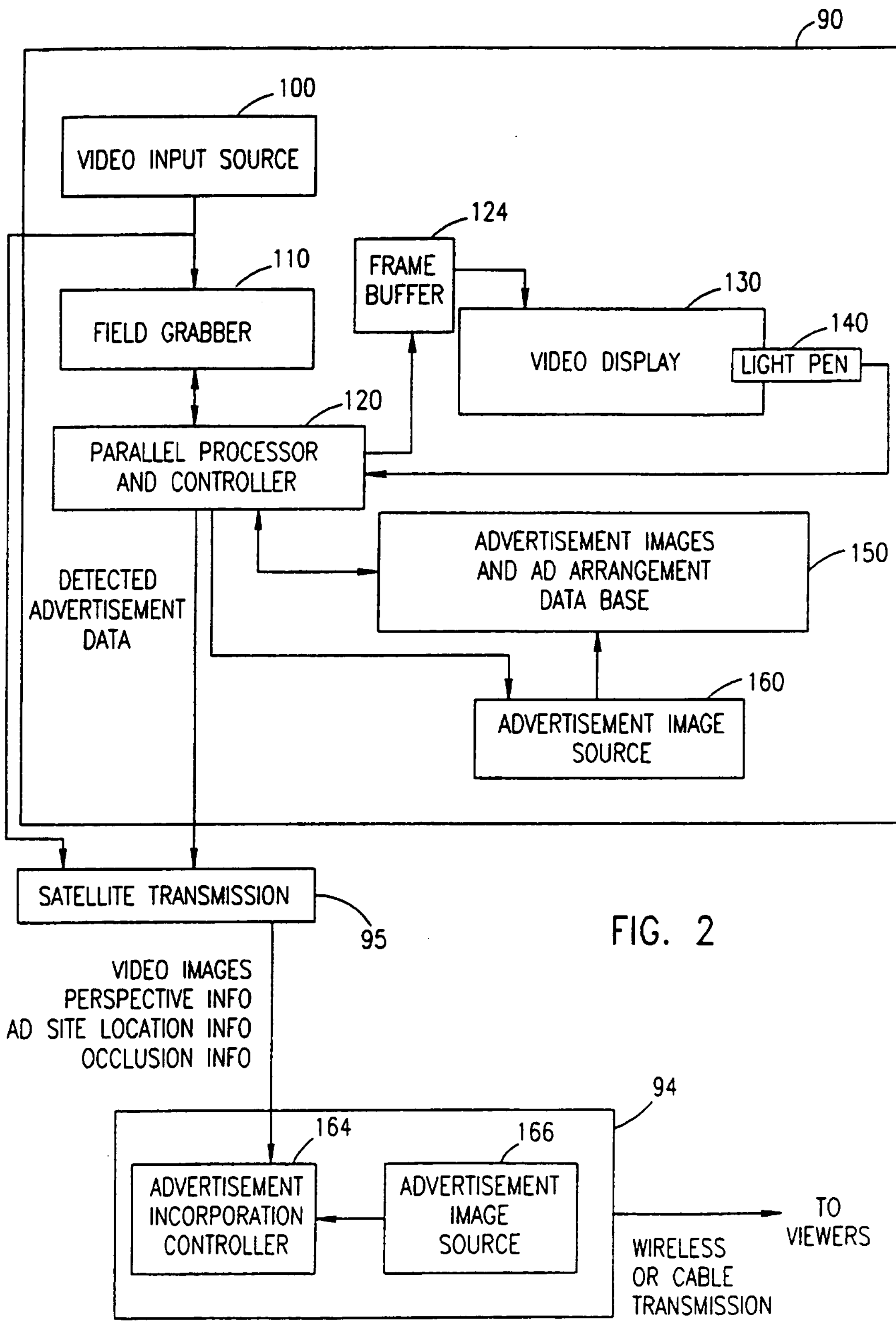

90
100
VIDEO INPUT SOURCE
124
FRAME BUFFER
110
FIELD GRABBER
130
VIDEO DISPLAY
140
LIGHT PEN
120
PARALLEL PROCESSOR AND CONTROLLER
ADVERTISEMENT IMAGES AND AD ARRANGEMENT DATA BASE
150
DETECTED ADVERTISEMENT DATA
160
ADVERTISEMENT IMAGE SOURCE
SATELLITE TRANSMISSION
95
VIDEO IMAGES
PERSPECTIVE INFO
AD SITE LOCATION INFO
OCCLUSION INFO
94
164
166
ADVERTISEMENT INCORPORATION CONTROLLER
ADVERTISEMENT IMAGE SOURCE
TO VIEWERS
WIRELESS OR CABLE TRANSMISSION

FIG. 2

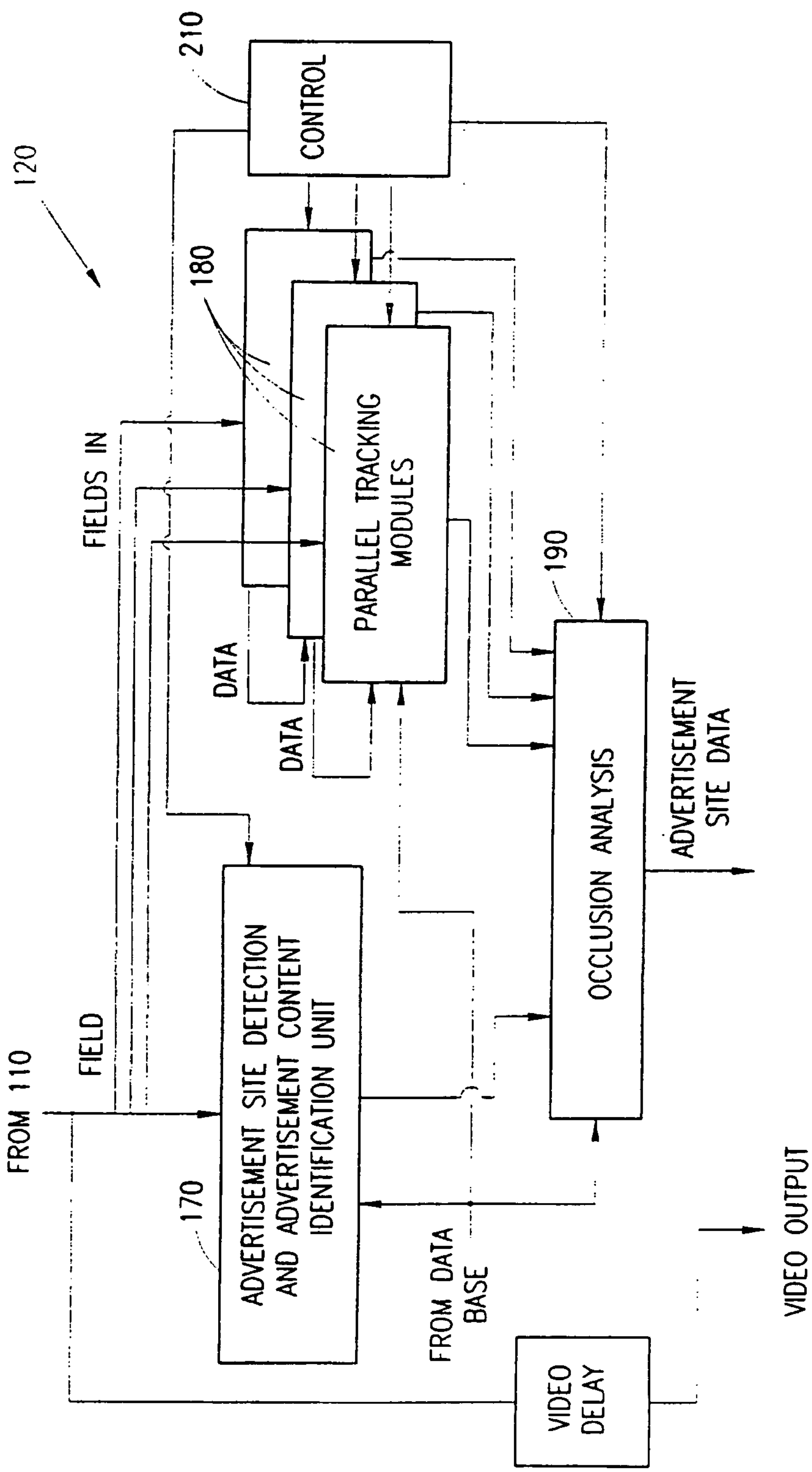

120
210
CONTROL
180
PARALLEL TRACKING MODULES
FIELDS IN
DATA
DATA
190
OCCLUSION ANALYSIS
ADVERTISEMENT SITE DATA
FROM 110
FIELD
ADVERTISEMENT SITE DETECTION AND ADVERTISEMENT CONTENT IDENTIFICATION UNIT
170
FROM DATA BASE
VIDEO DELAY
VIDEO OUTPUT

FIG. 3

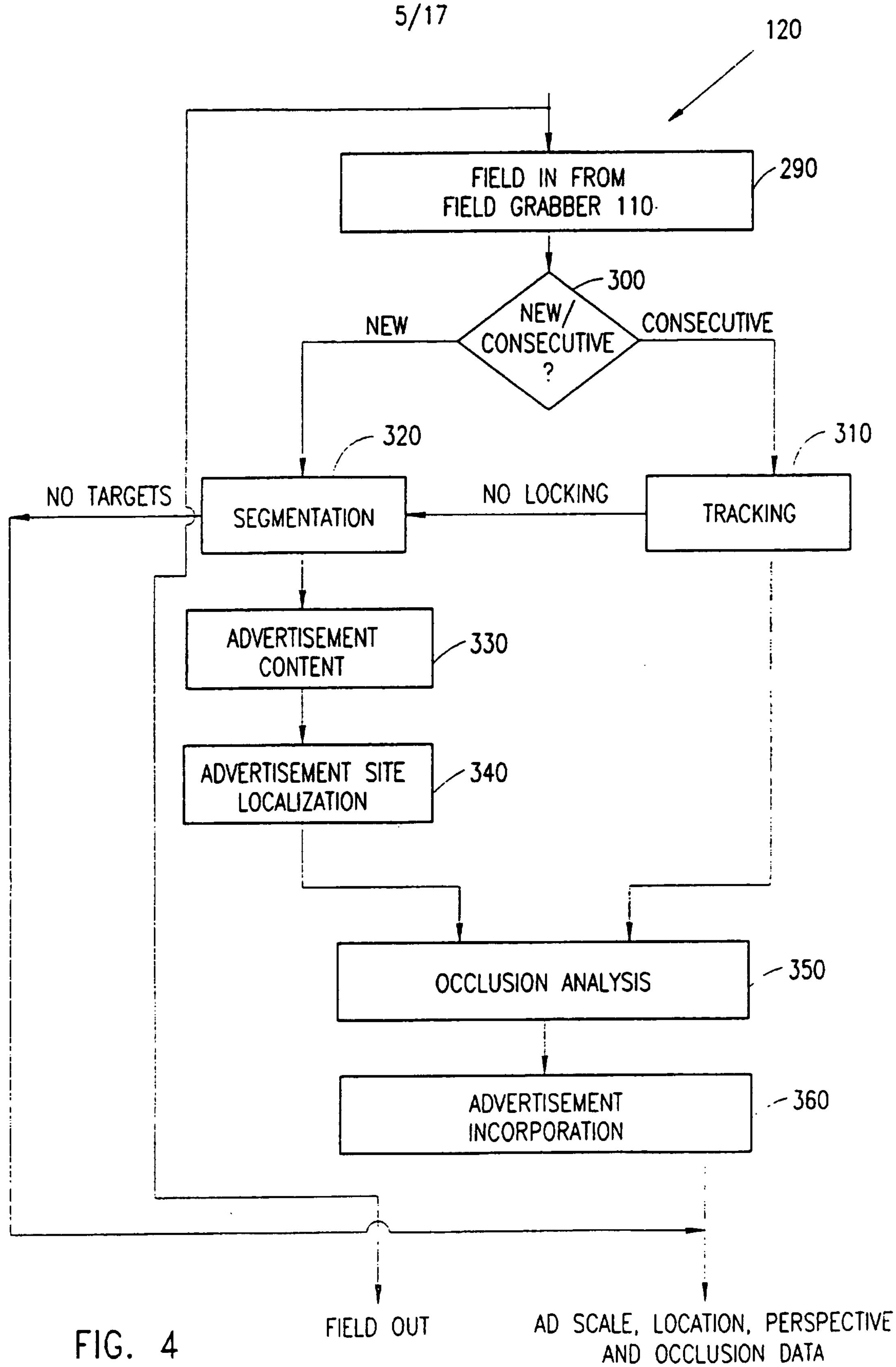

120
FIELD IN FROM FIELD GRABBER 110
290
NEW/ CONSECUTIVE ?
300
NEW
CONSECUTIVE
320
310
NO TARGETS
SEGMENTATION
NO LOCKING
TRACKING
ADVERTISEMENT CONTENT
330
ADVERTISEMENT SITE LOCALIZATION
340
OCCLUSION ANALYSIS
350
ADVERTISEMENT INCORPORATION
360
FIELD OUT
AD SCALE, LOCATION, PERSPECTIVE AND OCCLUSION DATA

FIG. 4

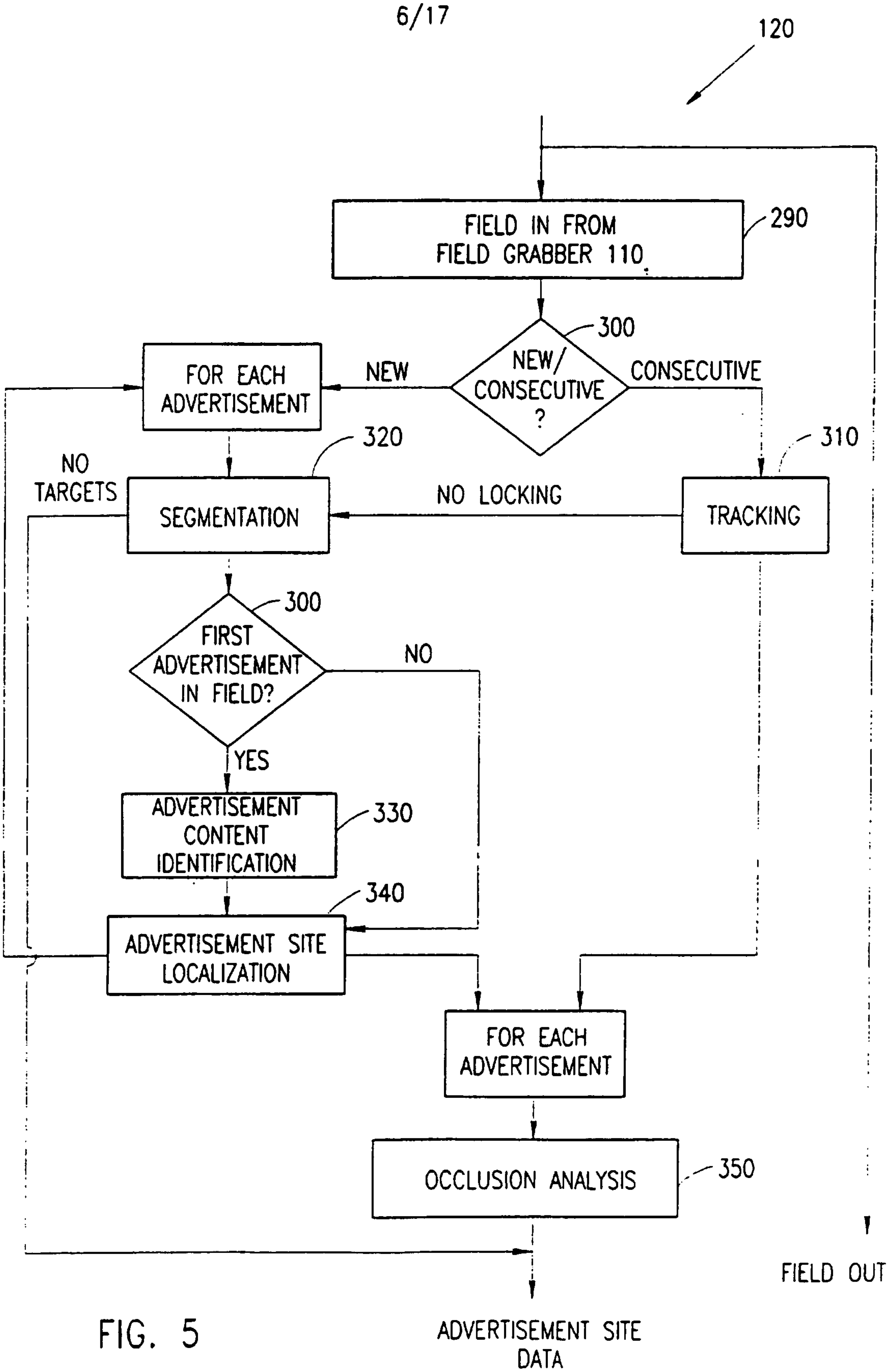

120
FIELD IN FROM FIELD GRABBER 110
290
300
NEW/ CONSECUTIVE ?
NEW
CONSECUTIVE
FOR EACH ADVERTISEMENT
320
310
NO TARGETS
SEGMENTATION
NO LOCKING
TRACKING
300
FIRST ADVERTISEMENT IN FIELD?
NO
YES
ADVERTISEMENT CONTENT IDENTIFICATION
330
340
ADVERTISEMENT SITE LOCALIZATION
FOR EACH ADVERTISEMENT
OCCLUSION ANALYSIS
350
FIELD OUT
ADVERTISEMENT SITE DATA

FIG. 5

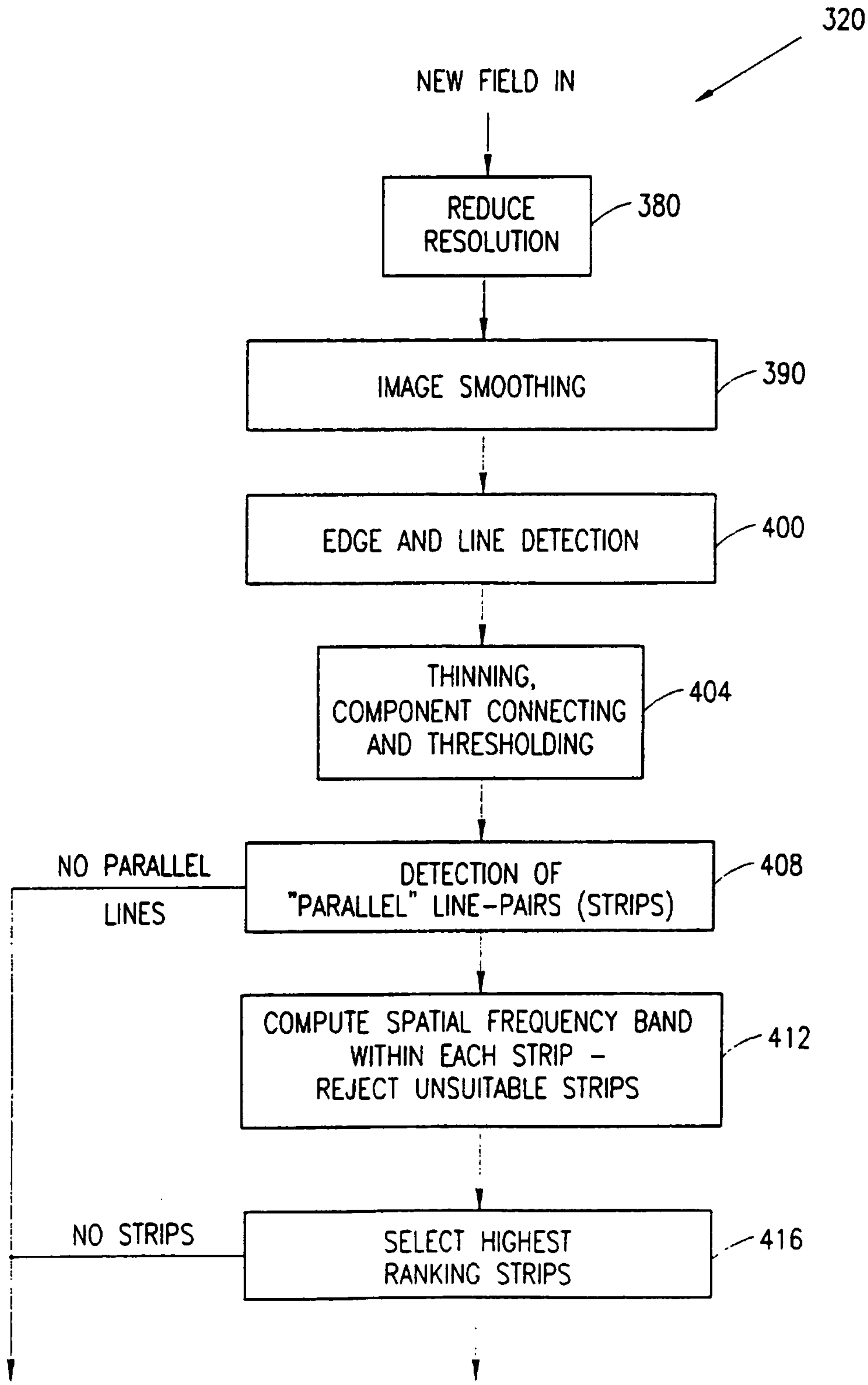

320
NEW FIELD IN
REDUCE RESOLUTION
380
IMAGE SMOOTHING
390
EDGE AND LINE DETECTION
400
THINNING, COMPONENT CONNECTING AND THRESHOLDING
404
NO PARALLEL LINES
DETECTION OF "PARALLEL" LINE-PAIRS (STRIPS)
408
COMPUTE SPATIAL FREQUENCY BAND WITHIN EACH STRIP - REJECT UNSUITABLE STRIPS
412
NO STRIPS
SELECT HIGHEST RANKING STRIPS
416

FIG. 6

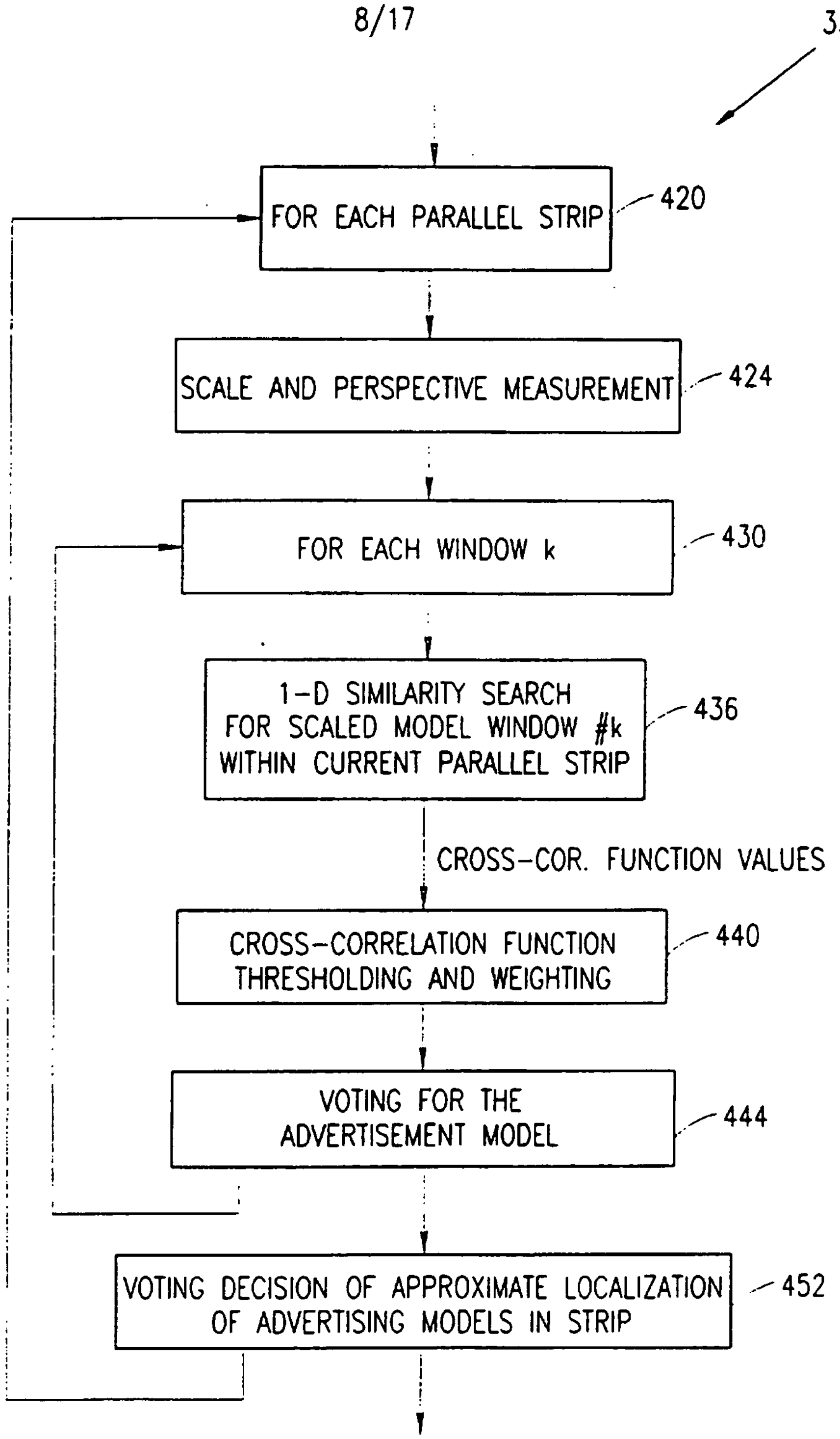

330
FOR EACH PARALLEL STRIP
420
SCALE AND PERSPECTIVE MEASUREMENT
424
FOR EACH WINDOW k
430
1-D SIMILARITY SEARCH FOR SCALED MODEL WINDOW #k WITHIN CURRENT PARALLEL STRIP
436
CROSS-COR. FUNCTION VALUES
CROSS-CORRELATION FUNCTION THRESHOLDING AND WEIGHTING
440
VOTING FOR THE ADVERTISEMENT MODEL
444
VOTING DECISION OF APPROXIMATE LOCALIZATION OF ADVERTISING MODELS IN STRIP
452

FIG. 7

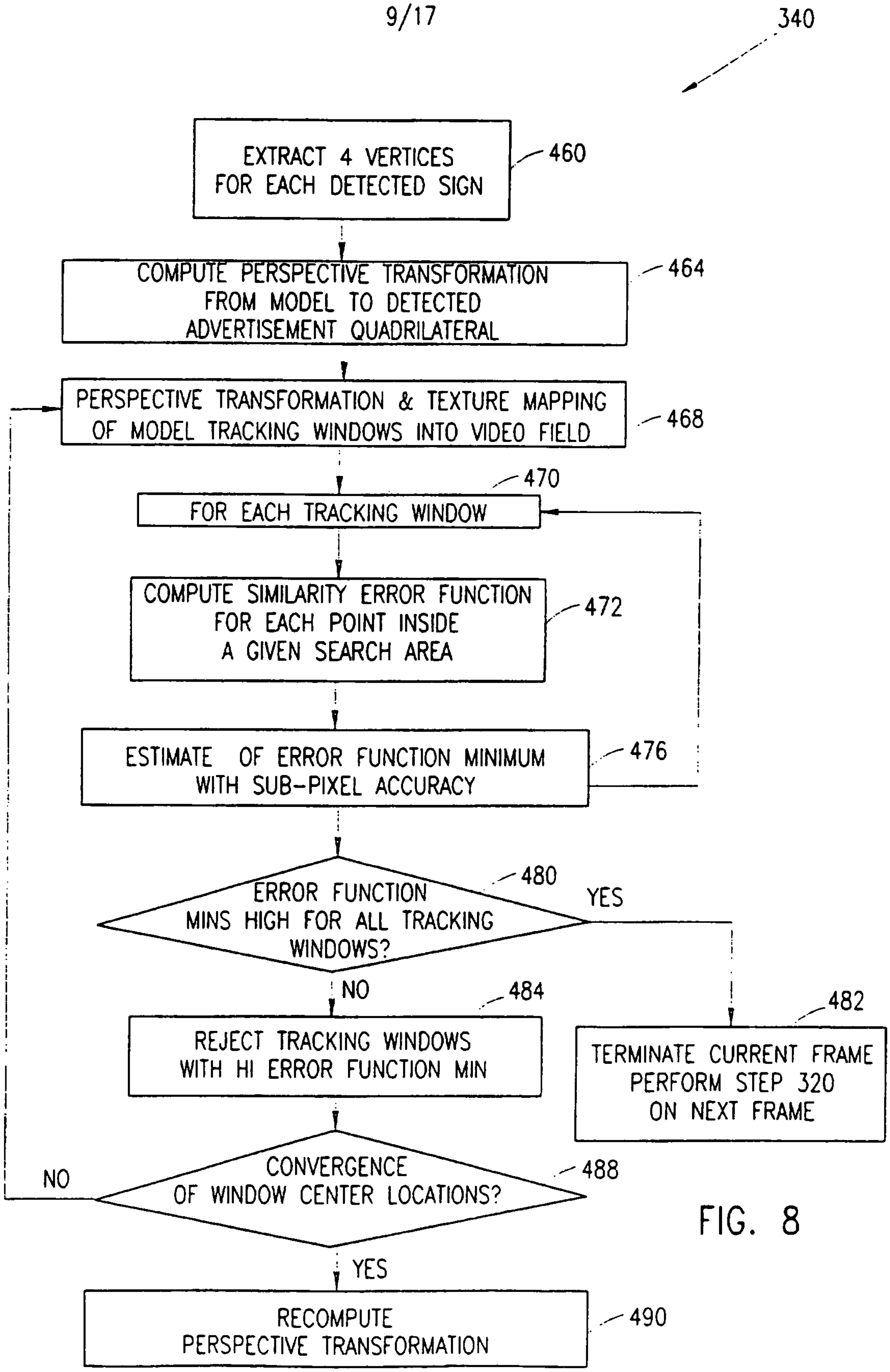

340
EXTRACT 4 VERTICES FOR EACH DETECTED SIGN
460
COMPUTE PERSPECTIVE TRANSFORMATION FROM MODEL TO DETECTED ADVERTISEMENT QUADRILATERAL
464
PERSPECTIVE TRANSFORMATION & TEXTURE MAPPING OF MODEL TRACKING WINDOWS INTO VIDEO FIELD
468
FOR EACH TRACKING WINDOW
470
COMPUTE SIMILARITY ERROR FUNCTION FOR EACH POINT INSIDE A GIVEN SEARCH AREA
472
ESTIMATE OF ERROR FUNCTION MINIMUM WITH SUB-PIXEL ACCURACY
476
ERROR FUNCTION MINS HIGH FOR ALL TRACKING WINDOWS?
480
YES
NO
REJECT TRACKING WINDOWS WITH HI ERROR FUNCTION MIN
484
TERMINATE CURRENT FRAME PERFORM STEP 320 ON NEXT FRAME
482
CONVERGENCE OF WINDOW CENTER LOCATIONS?
488
NO
YES
RECOMPUTE PERSPECTIVE TRANSFORMATION
490

FIG. 8

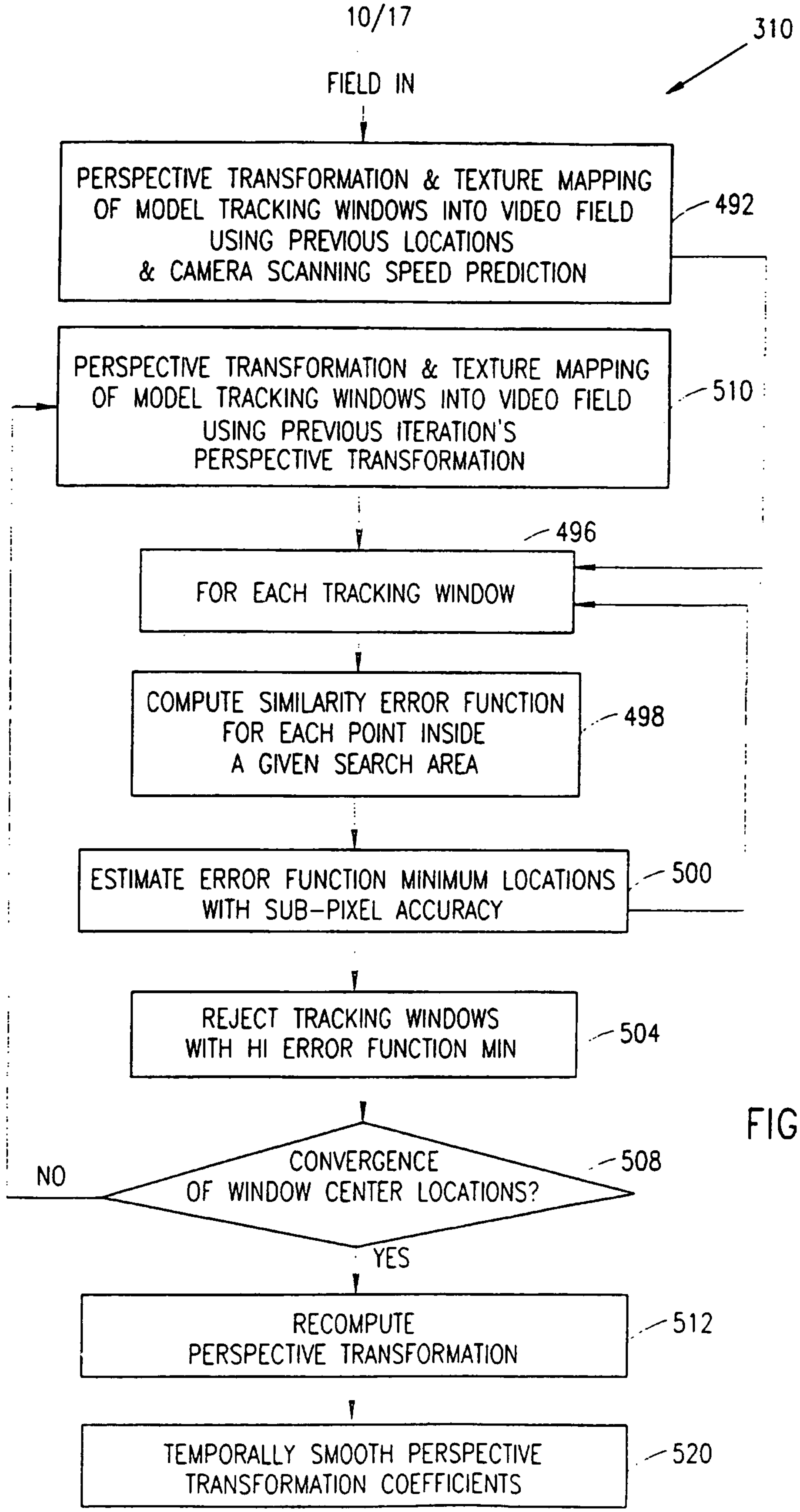
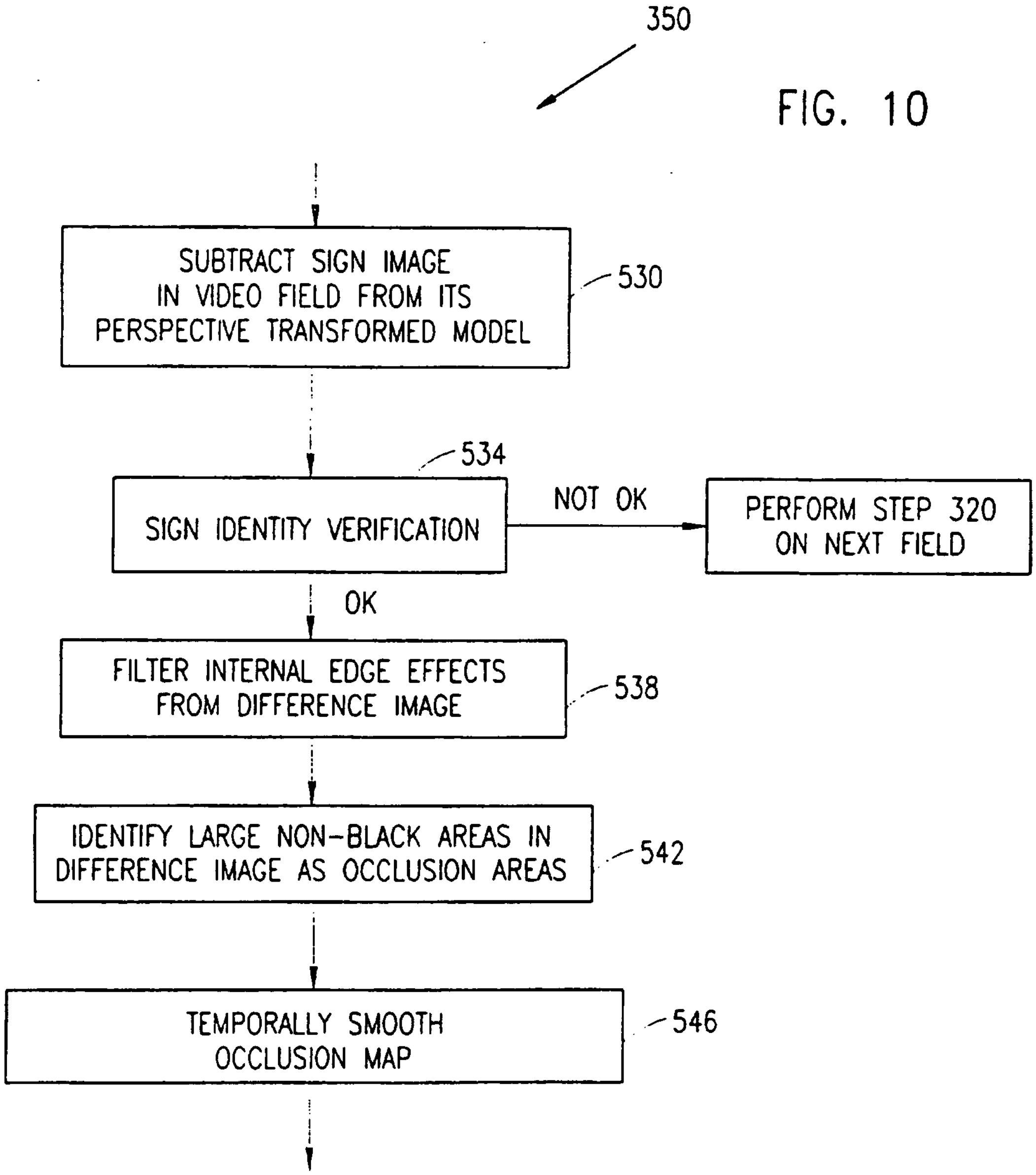

310
FIELD IN
PERSPECTIVE TRANSFORMATION & TEXTURE MAPPING OF MODEL TRACKING WINDOWS INTO VIDEO FIELD USING PREVIOUS LOCATIONS & CAMERA SCANNING SPEED PREDICTION
492
PERSPECTIVE TRANSFORMATION & TEXTURE MAPPING OF MODEL TRACKING WINDOWS INTO VIDEO FIELD USING PREVIOUS ITERATION'S PERSPECTIVE TRANSFORMATION
510
496
FOR EACH TRACKING WINDOW
COMPUTE SIMILARITY ERROR FUNCTION FOR EACH POINT INSIDE A GIVEN SEARCH AREA
498
ESTIMATE ERROR FUNCTION MINIMUM LOCATIONS WITH SUB-PIXEL ACCURACY
500
REJECT TRACKING WINDOWS WITH HI ERROR FUNCTION MIN
504
CONVERGENCE OF WINDOW CENTER LOCATIONS?
508
NO
YES
RECOMPUTE PERSPECTIVE TRANSFORMATION
512
TEMPORALLY SMOOTH PERSPECTIVE TRANSFORMATION COEFFICIENTS
520

FIG. 9

FIG. 10


350
SUBTRACT SIGN IMAGE IN VIDEO FIELD FROM ITS PERSPECTIVE TRANSFORMED MODEL
530
534
SIGN IDENTITY VERIFICATION
NOT OK
PERFORM STEP 320 ON NEXT FIELD
OK
FILTER INTERNAL EDGE EFFECTS FROM DIFFERENCE IMAGE
538
IDENTIFY LARGE NON-BLACK AREAS IN DIFFERENCE IMAGE AS OCCLUSION AREAS
542
TEMPORALLY SMOOTH OCCLUSION MAP
546

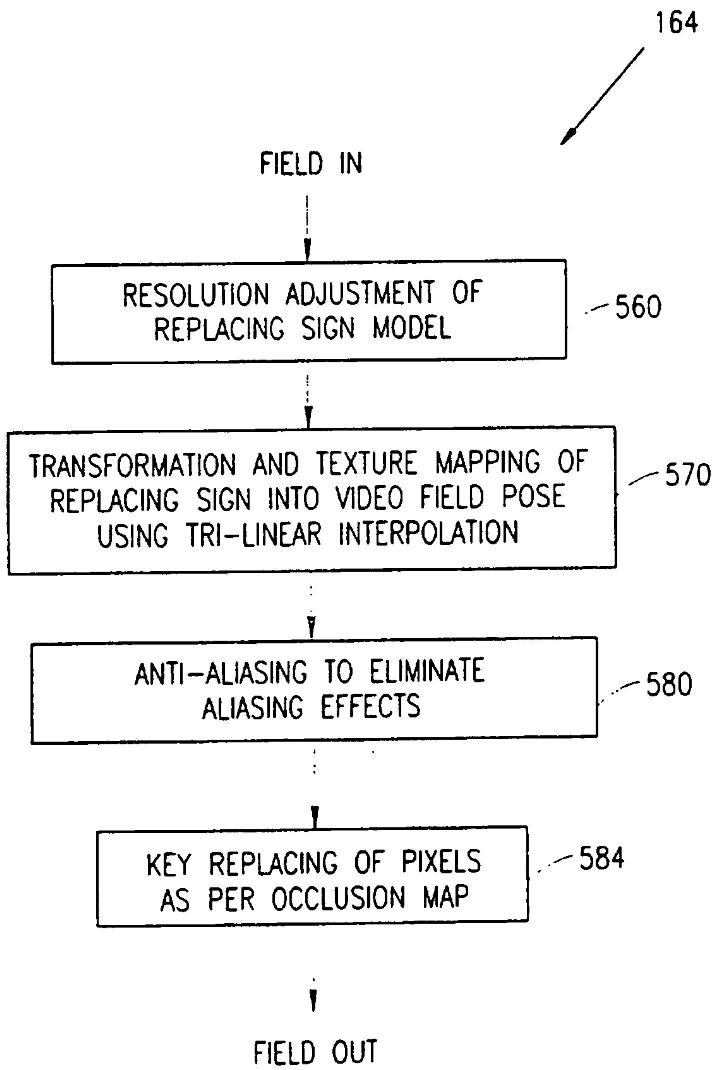

164
FIELD IN
RESOLUTION ADJUSTMENT OF REPLACING SIGN MODEL
560
TRANSFORMATION AND TEXTURE MAPPING OF REPLACING SIGN INTO VIDEO FIELD POSE USING TRI-LINEAR INTERPOLATION
570
ANTI-ALIASING TO ELIMINATE ALIASING EFFECTS
580
KEY REPLACING OF PIXELS AS PER OCCLUSION MAP
584
FIELD OUT

FIG. 11

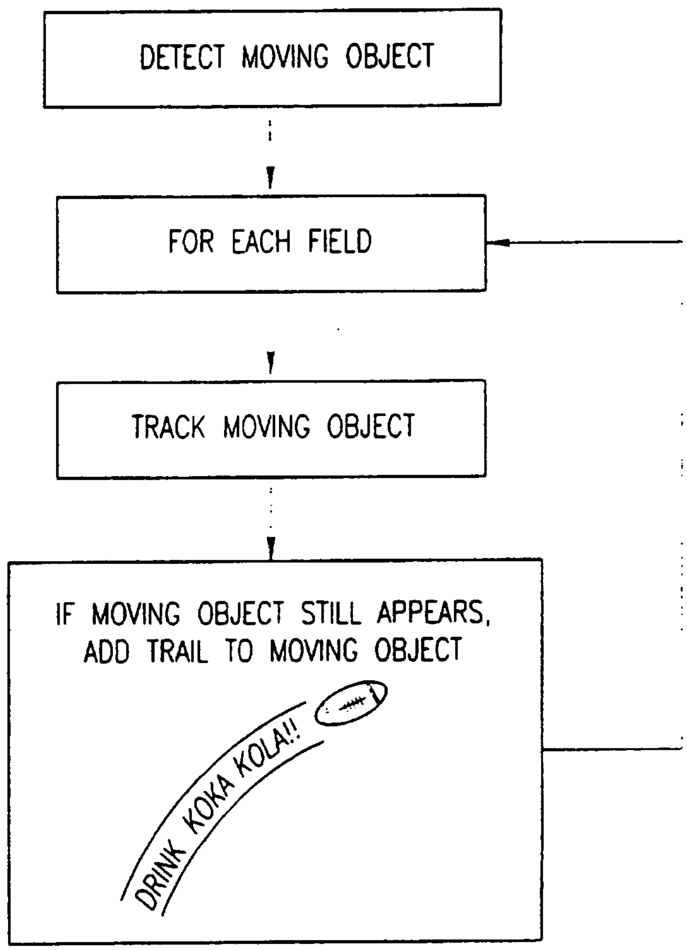

DETECT MOVING OBJECT
FOR EACH FIELD
TRACK MOVING OBJECT
IF MOVING OBJECT STILL APPEARS, ADD TRAIL TO MOVING OBJECT
DRINK KOKA KOLA!!

FIG. 12

FIG. 13

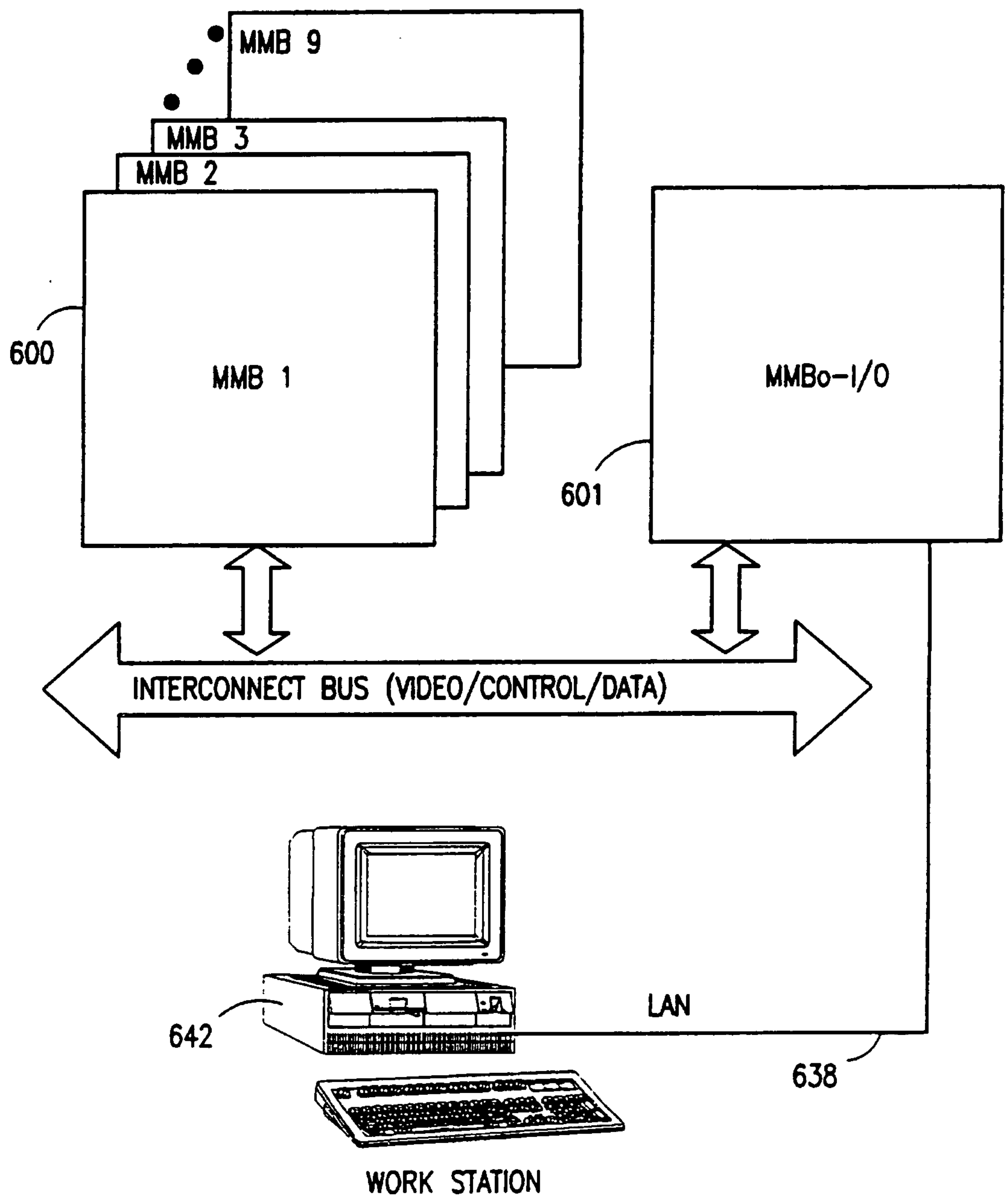

MMB 9
MMB 3
MMB 2
600
MMB 1
MMBo-I/O
601
INTERCONNECT BUS (VIDEO/CONTROL/DATA)
LAN
642
638
WORK STATION

FIG. 14

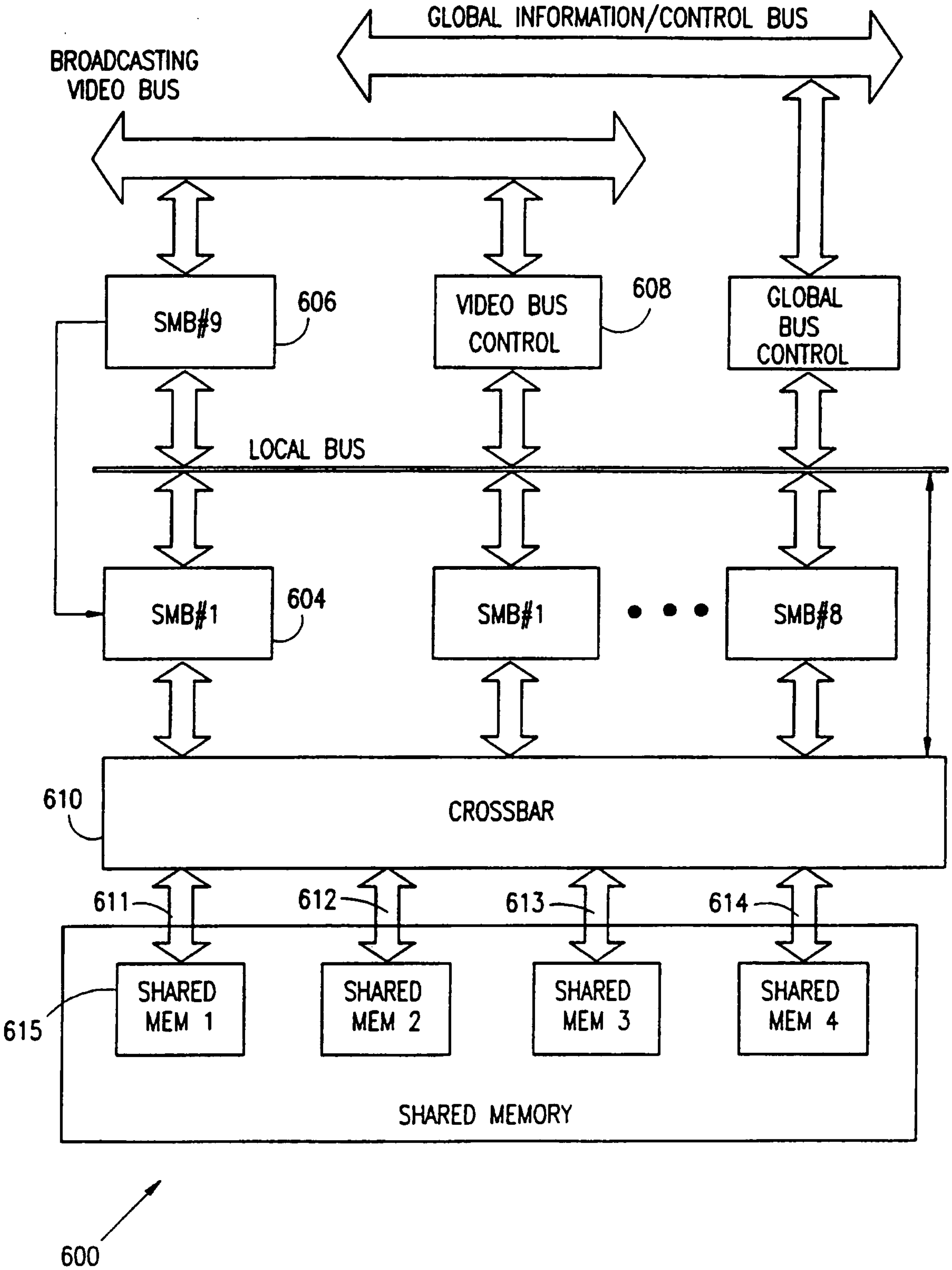

GLOBAL INFORMATION/CONTROL BUS
BROADCASTING VIDEO BUS
SMB#9
606
VIDEO BUS CONTROL
608
GLOBAL BUS CONTROL
LOCAL BUS
SMB#1
604
SMB#1
SMB#8
610
CROSSBAR
611
612
613
614
615
SHARED MEM 1
SHARED MEM 2
SHARED MEM 3
SHARED MEM 4
SHARED MEMORY
600

FIG. 15

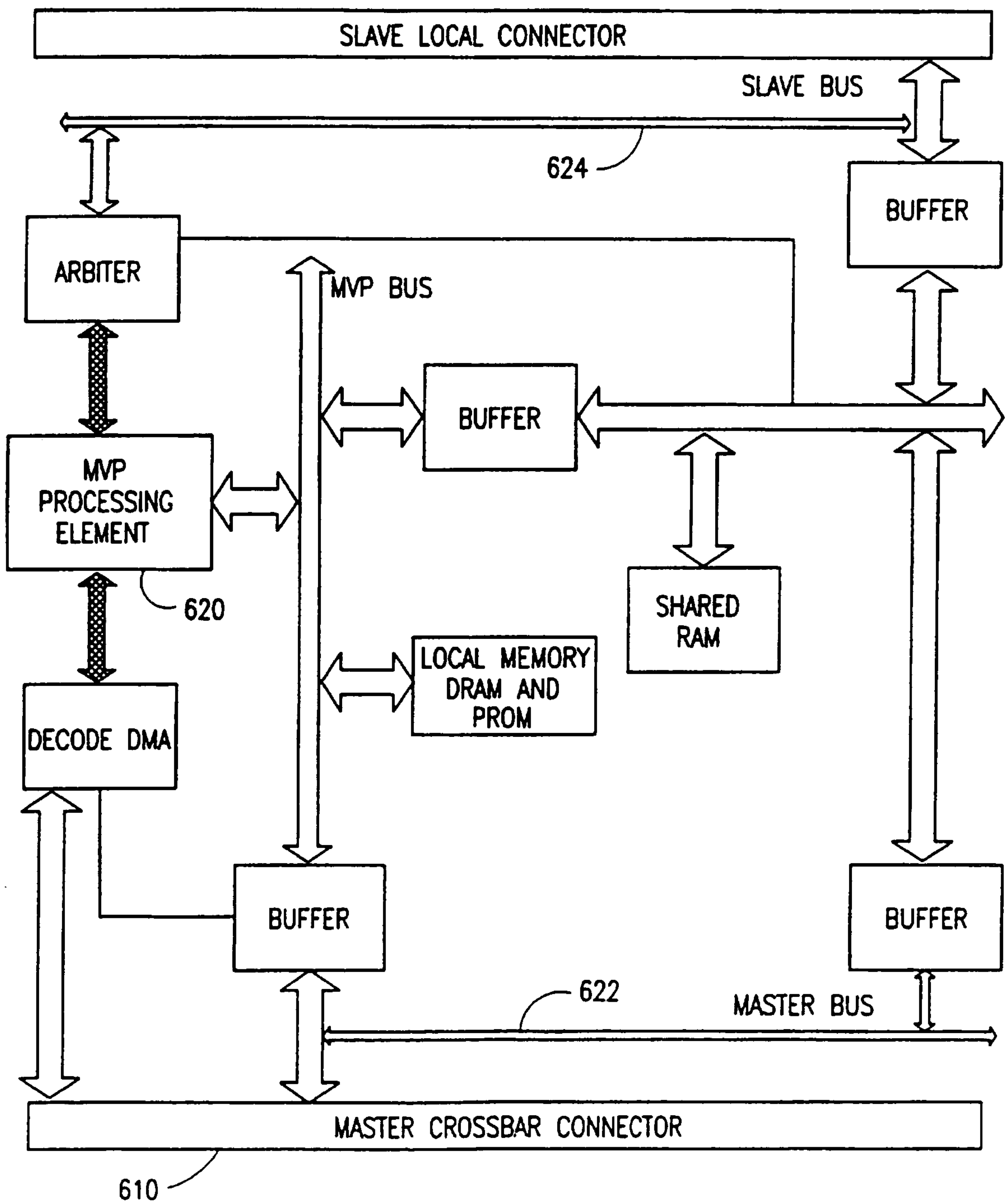
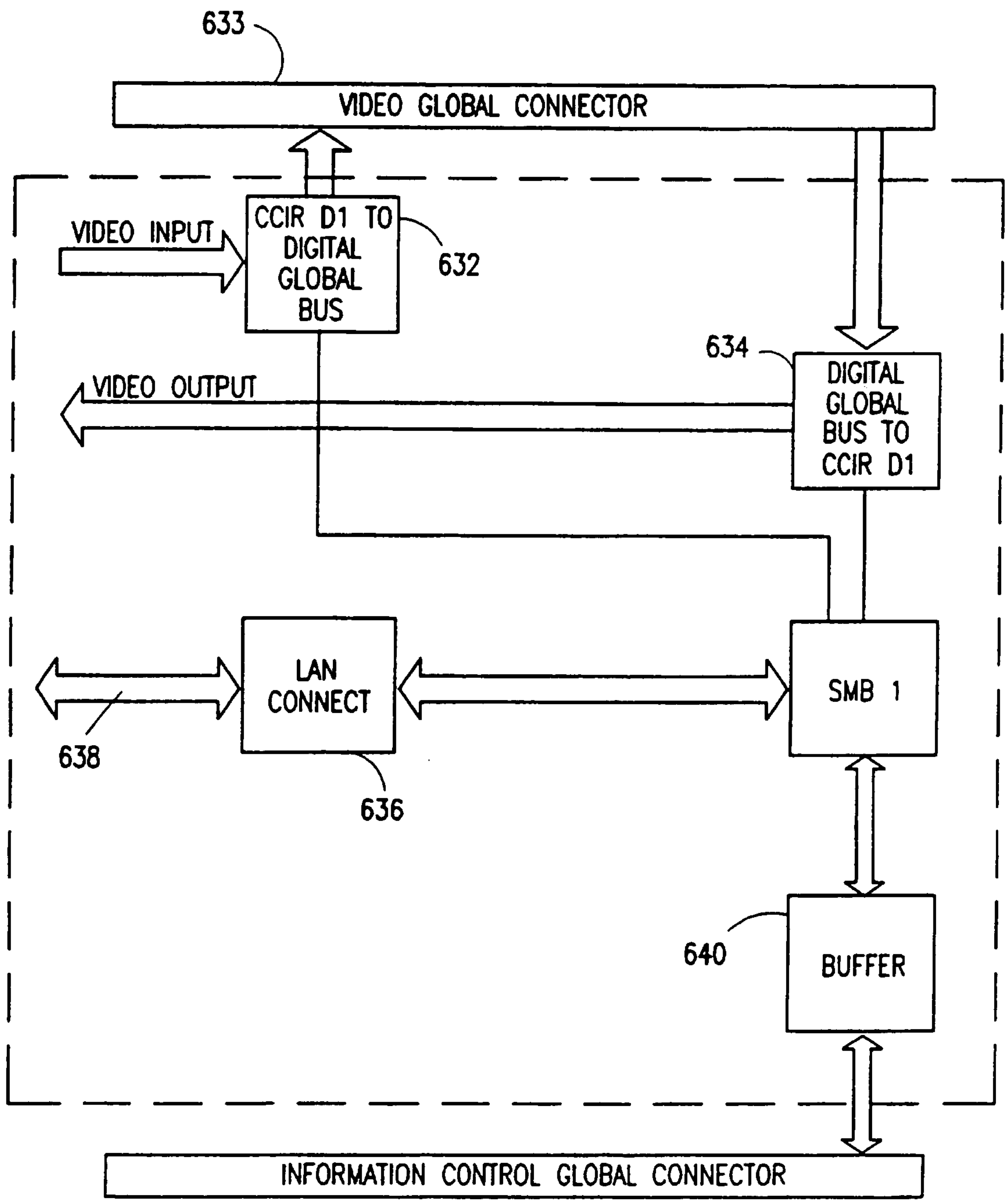

SLAVE LOCAL CONNECTOR
SLAVE BUS
624
BUFFER
ARBITER
MVP BUS
BUFFER
MVP
PROCESSING
ELEMENT
620
SHARED
RAM
LOCAL MEMORY
DRAM AND
PROM
DECODE DMA
BUFFER
BUFFER
622
MASTER BUS
MASTER CROSSBAR CONNECTOR
610

FIG. 16


633
VIDEO GLOBAL CONNECTOR
VIDEO INPUT
CCIR D1 TO DIGITAL GLOBAL BUS
632
634
VIDEO OUTPUT
DIGITAL GLOBAL BUS TO CCIR D1
LAN CONNECT
SMB 1
638
636
640
BUFFER
INFORMATION CONTROL GLOBAL CONNECTOR

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04N 5/262, 5/272, 5/222	A3	(11) International Publication Number: WO 96/05689 (43) International Publication Date: 22 February 1996 (22.02.96)

(21) International Application Number: PCT/US95/09946

(22) International Filing Date: 4 August 1995 (04.08.95)

(30) Priority Data:
110573 4 August 1994 (04.08.94) IL

(71) Applicant *(for all designated States except US)*: ORAD, INC. [US/US]; c/o Law Offices of Morse Geller, Suite 202, 116-16 Queens Boulevard, Forest Hills, NY 11375 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: TAMIR, Michael [IL/IL]; 13 Beit Zuri Street, 69122 Tel Aviv (IL). SHARIR, Avi [IL/IL]; 21 Ani Ma'amin Street, 67727 Ramat Hasharon (IL).

(74) Agents: GALLOWAY, Peter, D.; Ladas & Parry, 26 West 61st Street, New York, NY 10023 (US) et al.

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

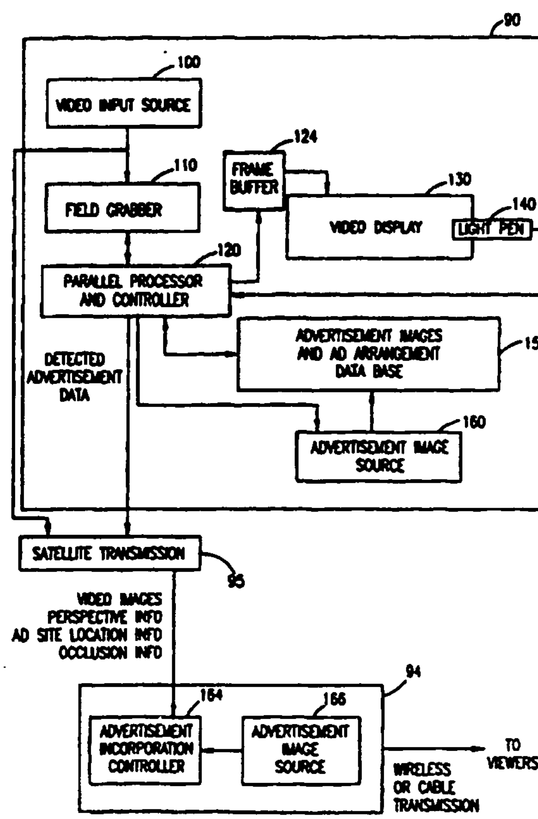
Published
With international search report.

(88) Date of publication of the international search report: 6 June 1996 (06.06.96)

(54) Title: APPARATUS AND METHOD FOR VIDEO BROADCASTING

(57) Abstract

This invention dislcoses an apparatus (90) for replacing a portion of each of a sequence of existing images with a new image, the apparatus comprises a frame grabber (110) operative to grab a sequence of frames respectively representing the sequence of existing images, a localizer (120) operative to detect at least one site within each existing image at which the new image is to be incorporated, a perspective transformer (120) operative to detect the perspective at which the site is imaged and a transmitter (95) operative to transmit to each of a plurality of remote locations (94), for each frame the existing image represented in the frame, the coordinates of the site, and the perspective at which the site is imaged.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/09946

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :H04N 5/262, 5/272, 5/222
US CL :348/722, 461, 9, 588, 589, 600
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

none

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

none

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,264,933 (ROSSER et al.) 23 November 1993, see the entire document.	1-10
Y,P	US, A, 5,436,672 (MEDIONI ET AL.) 25 July 1995, see abstract and FIG. 1.	1-10
A	US, A, 5,029,014 (LINDSTROM) 02 July 1991, see FIG.1.	1, 2
A	US, A, 5,424,770 (SCHMELZER ET AL.) 13 June 1995, see abstract and column 3, line 40 through column 6, line 29.	1-10
A	US, A, 5,200,825 (PERINE) 06 April 1993, FIG. 1 and column 2, line 37 to column 3, line 52.	1-10

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 NOVEMBER 1995	26 JAN 1996

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	SAFET METJAHIC
Facsimile No. (703) 305-3230	Telephone No. (703) 305-4741

Form PCT/ISA/210 (second sheet)(July 1992)☆

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 5,283,639 (ESCH ET AL.) 01 February 1994, column 1, line 42 to column 2, line 68.	1-10
A,P	US, A, 5,355,450 (GARMON ET AL.) 11 October 1994, column 9, line 19 to column 11, line 36.	1-10
A	US, A, 4,974,085 (CAMPBELL ET AL.) 27 November 1990, FIG. 1.	1-10

BEST AVAILABLE COPY

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/09946

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-10

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

348/720-722, 461, 9, 588, 589, 598-600, 705, 586, 584, 578, 576, 571, 552, 473
IPC(6): H04N 5/262, 5/272, 5/222, 5/265, 5/268, 5/278

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-10, drawn to an apparatus for replacing a portion of each of a sequence of existing images with a new image, which is classified in class 348, subclass 584.
Group II, claims 11-17, drawn to a real-time video processor and a method for controlling pipelined performance of a multi-task video processor, which is classified in class 348, subclass 571.
The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The claims of Group I describe an apparatus and method for replacing one image or a portion of one image with another image whereas the claims of group II describe an apparatus and method for controlling flow of image data over a data bus. The inventions of the two groups do not share the same special technical features and they are directed to different inventive concepts.